Contents

Chapter 1: Introduction to Thermodynamics

Chapter 2: Waves and Oscillations

Chapter 4: Diffraction

Chapter 5: Polarisation

Chapter 7: Laser

Unit - I

Introduction to Thermodynamics: Fundamental Ideas of Thermodynamics, The Continuum model, The Concept of a "System", "State", "Equilibrium", "Process", Equations of State, Heat, Zeroth law of Thermodynamics, Work, first and second laws of thermodynamics, entropy.

Chapter 1

Introduction to Thermodynamics

1.1 Introduction to Thermodynamics

Thermodynamics is the science that deals with the rules according to which bodies exchange energy. In the earlier times, thermodynamics was explained in terms of the relationship between mechanical and heat energy. However, with time thermodynamics got more developed by taking into account different branches of physics and chemistry showing inter-relationships between heat and all other forms of energy such as magnetic, chemical, electrical etc. We all know that all these forms of energy are inter-convertible thereby following the law of conservation of energy. The entire structure of thermodynamics rests on two well-known laws: the first and the second law of thermodynamics. The first law of thermodynamics is based on the law of conservation of energy i.e., energy can neither be created nor be destroyed but can be converted from one form to the other. The second law of thermodynamics expounds the idea that it is impossible to convert a given amount of heat fully into work. It explains the conditions under which conversion of heat into work or vice a versa can take place and it gives also an idea about the direction of heat transfer. The second law is also known as the law of increasing entropy.

1.2 The Continuum Model

A continuum model assumes that the substance of the object fills the space it occupies. Modelling objects in this way ignores the fact that matter is made of atoms, and so is not continuous; however, on length scales much greater than that of inter-atomic distances, such models are highly accurate. These models can be used to derive differential equations that describe the behaviour of such objects using physical laws, such as mass conservation, momentum conservation, and energy conservation, and some information about the material is provided by constitutive relationships. Continuum mechanics deals with the physical properties of solids and fluids which are independent of any particular coordinate system in which they are observed.

As a consequence of continuum approach fluid properties such as density, viscosity, thermal conductivity, temperature etc can be considered as continuous function of space and times.

1.3 Thermodynamic System: State Variables

In the thermodynamics, internal state of the system is our main concern and it is usual to describe it in terms of macroscopic quantities i.e., large scale or bulk properties of the system which depends on the internal state. These macroscopic quantities are called thermodynamic coordinates or state variables and the system is known as thermodynamic system. In case of gaseous system, the thermodynamic coordinates used are pressure(P), Volume(V), Temperature (T) and Entropy(S). Ex:

(a)a stretched bar has thermodynamic coordinates as length of the bar, tension to which it is subjected and the temperature(b) thin films like oil films on water or soap bubbles behave as stretched membranes and the thermodynamic coordinates are area of the film, surface tension and temperature. In these examples, we have taken pressure to be constant (atmospheric pressure) and we consider very small changes in volume.
An equation connecting the thermodynamic coordinates of the system (one of which becomes a dependent variable in the process) is called the Equation of state of the system and expresses the individual behavior of the system.

1.4 Zeroth Law of Thermodynamics: Concept of Temperature

This law was formulated after the first and the second law of thermodynamics and explains the concept of temperature of a system. Temperature can be defined as degree of hotness or coldness of a body.
It is seen that whenever a hot body is placed near to cold body, a change in temperature is observed. In other words, the body at higher temperature always gets cooler and the body with the lower temperature gets hotter. Thus, we can say that an energy or a heat exchange takes place between the two. As a result, the temperature of the bodies become equal and no more heat exchange takes place and we then can say that the two bodies are in thermal equilibrium. The thermal energy of each body now remains constant. If the two systems are not in thermal equilibrium, they will be at different temperatures. Thus, the temperature is not an absolute term rather it is a relative term and can now be defined as a quantity indicating the direction of heat exchange.
This concept of thermal equilibrium further led to the zeroth law of thermodynamics which states that "Two bodies A and B, each in thermal equilibrium with a third body C, are in thermal equilibrium with each other". Similarly, if we take a number of systems which are in thermal equilibrium with each other, the common property of the system can be represented by a single numerical value i.e., temperature. Zeroth law of thermodynamics thus indicates that the necessary and sufficient condition for thermal equilibrium is equality of temperature.
According to kinetic theory of gases, temperature is a measure of the average kinetic energy of the translational motion of the molecules of an ideal gas and signifies that the temperature of body is closely related to the energy of motion of its molecules. The greater the average kinetic energy per molecule of a body, the higher is its temperature. It means to change the temperature of a body, the average kinetic energy per molecule of the body is to be change. The best mechanism for this is, by heating or by cooling. For heating a body, heat energy must be supplied to it and to cool the body the heat energy must be taken away from it.

1.5 Heat

Julius Robert Mayer (1814-1878), a German doctor serving as a physician on a ship, conceived on 1842 the idea of the equivalence of heat and work. He hypothesized that heat is a form of energy. However, the relationship between heat and work was established in 1850 experimentally by James Joule. 1818-1989 the British physicist conducted a long series of experiments that conclusively demonstrated the equivalence of mechanical energy and heat energy.
Heat is basically the energy transferred or energy transferred from a hot body to the cold body. For example: If we consider a pan containing hot water (system), then after sometime the hot water starts getting colder as its temperature starts decreasing and tends to approach the room temperature. This is all because of exchange of energy between water (the system) and the surroundings. We now define

heat in a more general way. Heat is the energy that flows from one body or the system to another solely as a result of the temperature difference between them.
Thus, heat is not an intrinsic property of a body, a body containing certain amount of heat is meaningless. Also, if at any point we say the flow of heat stops, the word heat becomes meaningless. The term therefore can be used only and only when there is transfer of energy between two or more systems.

1.6 Heat and Work

The conversion of mechanical energy into heat and the reverse process of obtaining mechanical work at the expense of heat are the greatest interest in engineering. For example, in a thermal power station, thermal energy in the form of steam at high temperature and pressure drives the turbine and is converted into the mechanical energy. The turbine in turn drives the rotor of a generator which produces electricity.
Thermodynamics is the science that deals with work and heat, and those properties of substances related to heat and work. Many of the properties of the substance and many phenomena can be studied either with a microscopic or a macroscopic point of view. For instance, let us consider the familiar example of monoatomic gas confined in a small container. Quantities characteristics of the atoms such as an energy of an atom, its velocity, its mass, etc. can be used to describe the properties of the gas. These quantities can be computed from the theoretical analysis but cannot be measured. All such quantities are said to be microscopic. We may also describe the properties of the gas using quantities whose value can be measured. Such quantities are called macroscopic quantities. Temperature, pressure and volume are examples of macroscopic quantities. These are the quantities the gas as a whole has, but have no meaning when applied to individual atoms. Thus, we cannot speak of the pressure of temperature of the molecule. It is possible to develop the principles of thermodynamics from a microscopic point of view. Historically, the central concepts of thermodynamics were developed from a macroscopic view point without reference to microscopic models and details of the structure of matter. We study here the principles of thermodynamics from the macroscopic point of view.

1.7 Thermodynamic Concepts

There are certain terms in thermodynamics which we need to understand:

1.7.1 System

A well-defined system is a system in which the principles of thermodynamics are usually stated or are well applicable. A system is defined in thermodynamics as a quantity of matter of fixed mass and identity. Let us take an example of a thermodynamic system of a quantity of gas enclosed in a cylinder fitted with a movable piston The system and its environment are distinctly defined by drawing a boundary between them. The system can interact with its environment, mainly in two ways - by transfer of heat or by doing work.

1.7.2 Thermodynamic Equilibrium

In mechanics, equilibrium means state of rest. In thermodynamics, the concept is somewhat broader. A system is said to be in thermodynamic equilibrium if none of the thermodynamic variables determining its state changes with time. Thermodynamic equilibrium is easily understood in the case of a monoatomic gas confined in a cylinder. If the temperature of a gas in the cylinder is same at all points in the cylinder and the temperature of the walls of the cylinder is also the same, then the gas is said to be in thermal equilibrium with the cylinder. The heat of the gas does not flow from one part of the cylinder to the other. Further, when neither the pressure nor the chemical composition of the gas changes, it is said to be in thermodynamic equilibrium. Thus, a system will be in a state of thermodynamic equilibrium if it satisfies all the conditions mechanical, thermal and chemical equilibrium. This can further be explained as:

1. **Mechanical equilibrium:** For a system to be in mechanical equilibrium, the force exerted by the system must remain constant and uniform and must be balanced by the external force on the system. There should be no unbalanced forces acting on a part or the whole of the system.
2. **Thermal equilibrium:** For a system to be in thermal equilibrium there should be no difference in the temperatures between the parts of the system or between the system and the surroundings.
3. **Chemical equilibrium:** For a system to be in chemical equilibrium there should be no chemical reaction within the system or change in the internal structure because of it and no movement of any chemical constituent from one part of the system to the other on account of solution or diffusion etc.

Thus, we can clearly now define the state of thermodynamic equilibrium as the state when all the above discussed types of thermal equilibrium are satisfied to the fullest. When a system undergoes the thermodynamical equilibrium keeping surroundings unchanged, no motion will take place and no work will be done. However, if any one of the above-described equilibrium conditions are violated or unsatisfied the system will said to be in non-equilibrium state. E.g., if there is an unbalanced force either in the interior of the system or between the system and the surroundings, the mechanical equilibrium will be disturbed and as a result turbulence, waves and eddies would be set in the system or there can be an accelerated motion of the system itself. These things can then lead to non-uniformity in temperature within the system or a temperature difference between the system and the surroundings. The system may then pass-through non-equilibrium states, which then cannot be described in terms of the system as a whole.
A cartesian coordinate system is used to plot sets of values of P, V and T to indicate equilibrium states of a system. Different points on the graph correspond to different equilibrium states of the gas. An isolated system always reaches a state of thermodynamic equilibrium in course of time but can never depart from it spontaneously.

1.7.3 Process

Any thermodynamical state of a system can always be defined only with the help of thermodynamical coordinates or state variables of the system. By changing the thermodynamical coordinates one can change the states of system. This change of state by changing the thermodynamical coordinates is called a process. Let us consider two states of a system i.e., state A and state B, the change of state from A to B or vice a versa is a process.
Some of the typical processes are:

1.**Isothermal process:** If the change in pressure and volume of gaseous system or in other words if a system is perfectly conducting in such a way that its temperature remains the same throughout, it is called isothermal process. A graph between pressure and volume of gas at constant temperature is called an isothermal.
Let us suppose, a cylinder fitted with a piston contains gas which is maintained at room temperature and under atmospheric pressure. Work is done on the gas when the piston is pushed down such that the internal energy increases and so the temperature. If the temperature has to be maintained

constant, then the extra heat must be conducted away to the surroundings. Similarly, if the compressed gas is allowed to expand and push the piston up a little i.e., some external work is done, the internal energy decreases and temperature decreases. Again, if the temperature has to be maintained constant, heat must be conducted to it from the surroundings.
Thus, in an isothermal process heat must be quickly conducted from the gas to the surroundings or vice a versa and this can be possible only when the containing vessel is a perfect conductor of heat.
If we want to ensure that the temperature of the system is constant, the withdrawal or supply of heat must keep pace with the rise and fall of temperature of the system. This means that the change in volume or pressure must be small as well as slow.
Thus, an ideal isothermal process must be infinitely slow and must consist of infinitely small steps in perfect thermal communication with the surroundings.
If the system is a perfect gas, the standard gas equation $PV = rT$ or $PV= RT$ will be applicable during an isothermal process where r and R are the gas constants per unit mass or per gram molecule of the gas respectively.
If we put it in the form $PV = rT = k$ or $PV = RT = k$, it gives the equation of the isothermal of a perfect gas at a temperature T.

2.**Adiabatic process:** During such type of process, the working substance is perfectly insulated from the surroundings. When work is done on the working substance, there is a rise in temperature because the external work done on the working substance increases its internal energy. When work is done by the working substance, it is done at the cost of its internal energy. As the system is perfectly insulated from the surroundings, there is fall in temperature.
Thus, during an adiabatic process, the working substance is perfectly insulated from the surroundings. All along the process, there is change in temperature. A curve between pressure and volume during an adiabatic process is called an adiabatic curve. e.g., Sudden bursting of a cycle tube is an adiabatic process. Also, in the example quoted in isothermal process above, if the cylinder and piston be of a perfectly insulating material and the gas in the cylinder be compressed it will rise in temperature since the heat developed cannot possibly escape out to the surroundings. If the gas is allowed to expand or to do external work, it will do at the expense of its own internal energy, since no heat can possibly enter the cylinder from outside and thus, the temperature of the gas will fall.
According to the first law of thermodynamics,

$$dQ = dU + dW$$

here since $dQ = 0$ so $dW = - dU$
This indicates that a decrease in internal energy takes place when for the system which is equal to the external work done by the system.
An adiabatic process is always accompanied by a change in the temperature of the system. Since no material is a perfect insulator of heat, some heat may escape from or enter into the system unless the process is quick and sudden. An adiabatic process is thus quick and sudden and takes place in thermal isolation from the surroundings.

3.**Isochoric process**: If the volume of the system is kept constant but pressure and temperature changing then the process is called an isochoric process. This is possible when the working substance is taken in a non-expanding chamber, the heat supplied will increase the pressure and temperature. The work done is zero here because there is no change in volume. The whole of the heat supplied increases the internal energy. The P-V or the indicator diagram of such a process will be a vertical line parallel to the pressure axis. As there is no change in the volume of the system, the work done is

$$\delta W = \int P dV = 0$$

Hence, if dQ be the heat supplied to the system, then

$$dQ = dU + dW$$

which gives $dQ = dU$

i.e., the whole of the heat supplied goes to increase the internal energy of the system

4. **Isobaric process**: If the working substance is taken in an expanding chamber kept at a constant pressure the process is called isobaric process. Here, both the temperature and volume change. The P-V or the indicator diagram of such a process will be a horizontal line parallel to the volume axis.

1.7.4 Quasistatic Process

We can remember from the previous discussion in section 1.6.2 about how a thermal equilibrium is achieved when the temperature of every part of the system is same with that to the surroundings. Therefore, there is a need in any process to describe the states of the system in terms of the same coordinates as those of the system itself. The external forces must not be varied to a very greater extent such that the unbalanced forces remain infinitesimal.

A quasi-static process is thus defined as the process in which the deviation from the thermodynamic equilibrium is infinitesimal and all the states through which the system passes during a quasi-static process can be considered as equilibrium states and may therefore be expressed in terms of thermodynamic coordinates of the system as a whole.

A quasi-static process is actually an ideal concept. Practically, only under rigorous conditions it can be satisfied. If somehow, we can slow down the process by proceeding step by step gradually, nearly quasi-static state can be achieved. e.g. if we decrease the pressure on a gas enclosed in a cylinder in small gradual steps by moving the piston up extremely slowly, the process of expansion may well be regarded as quasi-static. On the other hand, if we suddenly raise the piston up, the gas expands all at once. The process is therefore, far from quasi-static and at no stage in the process can the gas be regarded to be in an equilibrium state.

Work done in a quasistatic process – P-V or Indicator Diagrams

Work done in a quasi-static process such as for example changing the volume of a chemical or a gaseous system can be easily obtained with the help of what are called indicator diagrams or P-V diagrams. The changes in volume and pressure are due to the movement of piston in cylinder.

Let us discuss the cyclic and non-cyclic processes.

(i)**Non-Cyclic process**

Consider pressure along x axis, volume on y axis and let us suppose that the gas in the cylinder undergoes a change in pressure and volume at constant temperature. Let the pressure and volume changes gradually or quasi-statically along the curve AB in the P-V diagram or indicator diagram for the gas and represents the whole operation of expansion of the gas from its initial state at A to its final state at B.

Let the piston moves through an infinitesimal distance δx, an infinitesimal amount of work is done by the gas, given by

$\delta W = PA\delta x = P\delta V$

Where A is the cross-sectional area of the piston and δV the infinitesimal increase in volume.

This work is represented by the shaded area in Fig.1.0.

Thus, work done by the gas during the whole expansion from volume V_1 at A to volume V_2 at B is
$W = \int_A^B \delta W = \int_{V_1}^{V_2} P\delta V$ = the entire area GABH under the curve

If the gas were compressed along the same path from a volume V_2 at B to volume V_1 at A work done on the gas would be the same W numerically, but opposite in sign i.e., equal to -W. Conventionally, work done by the gas is positive and work done on it or absorbed by it as negative.

The area under the curve depends on different types of shapes with the same endpoints A and B. It is seen that work done by the gas depends not only on the initial and final states but also on the intermediate states it passes through i.e., on the path along which the change occurs. So work is a path function and not a point function like pressure or temperature.

Thus, $W = \int_A^B \delta W = \int_{V_1}^{V_2} P\delta V$ cannot be integrated to give $PV_2 - PV_1 = W_2 - W_1$ as δWor $P\delta V$ is an inexact differential.

(ii) **Cyclic process**

Let us suppose the gaseous system is subjected to a series of changes of pressure and volume such that its initial pressure and volume (represented by point A) change along the curve ACB and finally attain the values represented by point B. Let it now be subjected to a different series of pressures so that it comes back to its original state A by a different path BDA. The system is then said to perform a thermodynamic cycle of operations.

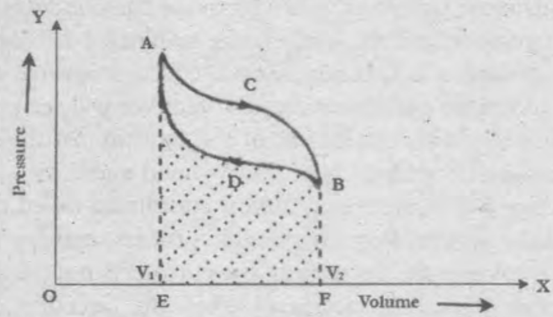


Fig. 1.0 Cyclic process

Work done by the system is given by the area ACBFE under curve ACB and shown shaded in Fig.1.0 The work done on the system is represented by area BDAEFB under the curve BDA and shown cross-shaded. So, the net external work(W) by the system is represented by the area of loop ACBDA and is equal to the difference of the areas ACBFEA and BDAEFB under the curves ACB and BDA respectively.

The area of loop is conventionally regarded as positive if the upperpart of the cycle is traversed towards the right i.e., expansion and negative, if it traverses towards left i.e., compression.

1.7.5 Reversible and Irreversible Processes

Reversible process

In thermodynamics, reversible process is the process in which for an infinitesimal small change in the external conditions can lead to changes taking place in the direct process but exactly repeated in the reverse order and in the opposite sense. The process should take place very slowly. There should be no loss of heat due to friction or radiation.

Let us consider a gas in a cylinder maintained at a certain temperature and pressure. This cylinder is attached with a frictionless piston. If the pressure is decreased, expansion of gas takes place very

slowly and constant temperature is maintained. The energy in this process is drawn continuously from the source(surroundings). If on the other hand pressure on the piston is increased, contraction of gas takes place slowly and temperature is maintained. The energy is liberated during compression and is given to sink(surroundings).
Hence, for a reversible process, there should be no loss of heat due to friction, radiation or conduction. If the changes take place very rapidly, the process will no more be reversible. Following are the conditions for a process to be reversible:
1.The pressure and the temperature of the working substance must not differ from those of the surroundings at any stage of the cycle of operation.
2.All the processes taking place in the cycle of operation must be infinitely slow.
3.There should be no heat losses.
Reversible process is only an ideal case. In reality there are always some losses of heat and the temperature and the pressure of the working substance changes from those to the surroundings.

Irreversible process
We know by this time that change of state from A to state B or vice a versa is a process. But the direction of the process is important and it always depends on a thermodynamical coordinate, called entropy. In practicality however all processes are actually not possible in the universe. e.g., Let us take two blocks 1 and 2 with different temperatures T_1 and T_2 such that $T_1 > T_2$ and they are in contact with each other but we insulate the system as whole from the surroundings. We will see that heat will be conducted between the two such that the temperature of block 1 falls and that of the temperature of block 2 rises and thermodynamical equilibrium is reached. Now reverse the direction i.e., the block 2 should transfer heat to 1 and initial conditions are restored. We will see that the reverse process will not be possible. In other words, the determination of the direction of the process cannot take place by knowing the thermodynamical coordinates in the two end states. So, in order to determine the direction, one has to introduce a new thermodynamical coordinate called the entropy of the system. Entropy is also the state of the system. For any possible process, entropy should increase or remain constant. The process which governs the decrease in the entropy is not possible.
Thus, a process is irreversible if entropy decreases when the direction of process is reversed. A process is said to be irreversible if it cannot be traced exactly back in the opposite direction. During an irreversible process, heat energy is always used to overcome friction. e.g.; all chemical reactions, all natural processes.

There are few more examples to understand the irreversibility of real thermal processes:
(i)**Free expansion of a gas:** Consider the example of a gas separated from vacuum by a membrane. If the membrane breaks, the gas fills the entire vessel. The system cannot be restored to the initial state unless the gas is compressed and the resulting heat is taken away. This, however, causes a change in the surroundings. Therefore, at the end of the process, the surroundings are not restored to their initial state. For this reason, the free expansion of a gas is an irreversible process.
(ii)**Diffusion:** Let us consider the case of two different gases, oxygen and hydrogen, separated by a membrane. If the membrane breaks, the diffusion of gases takes place spontaneously and a homogenous mixture of oxygen and hydrogen fill the entire volume. But the process will never reverse itself. The mixture of gases will never divide by itself in the initial components.
In practice, a mixture of gases can be divided into its initial components. But first, it involves application of energy. Secondly, the system will not pass again through the same intermediate states it

passes through the diffusion process. The system thus cannot be returned to its initial state without substantially altering the properties of the surroundings.

(iii) **Heat exchange:** Experience shows that the heat exchange like diffusion, is a one-way process. In heat exchange, energy is always transmitted from a body at the higher temperature to another body at lower temperature. Consequently, the heat exchange is always accompanied by an equalization of temperatures. The reverse process of transferring energy in the form of heat from cold bodies to hot bodies never occurs by itself. For instance, a hot cup of coffee cools through heat transfer to the surroundings. However, reverse will not happen.

1.7.6 Cycle

If a system in a given initial equilibrium state passes through a number of different changes of state and ultimately returns to its original equilibrium state, then the system is said to have gone through a thermodynamic cycle. The steps that constitute the cycle may be reversible or irreversible. If the system consists of a single homogenous substance and if all the steps are reversible, the cycle can be represented by a closed curve in a PV diagram as in Fig 1.1

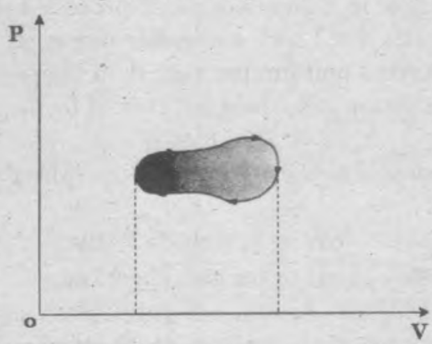


Fig.1.1 The PV diagram for an arbitrary cyclic process. The net work done in the process equals the area enclosed by the curve.

1.7.7 Heat Reservoir

A heat reservoir is a body from or to which heat can be transferred and it remains at constant temperature. For example: the atmosphere is a good heat reservoir. A heat reservoir from which heat is extracted by a body is called a high temperature heat reservoir (source) and heat reservoir to which heat is ejected from a body is called a low temperature reservoir(sink).

1.8 Work

The mechanical work performance by a force is most familiar form of work. In mechanics, work is said to have been performed when a force, acting on body displaces it through a distance x, the displacement being in the direction of force. That is:

$$W = F \times x \qquad (1)$$

The mechanical energy of a body changes when mechanical work is done on the body, Therefore, it serves as a measure of transfer of mechanical energy from one body to another. In general, we say work is a form of energy. Work is not stored in the body. Work is a form of transferring energy and is a measure of transferred energy.

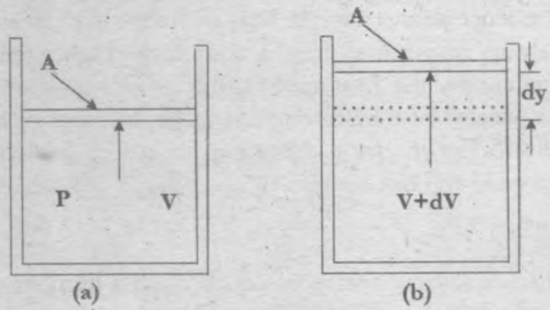


Fig. 1.2(a) Gas contained in a cylinder at a pressure P does work on the piston (b) As the piston moves the system expands from a volume V to a volume V + dV

Now let us examine the work done in a thermodynamic process. Consider a thermodynamic system such as gas contained in a cylinder fitted with a movable piston, as in Fig 1.2(a). In equilibrium the gas occupies a volume V and exerts a uniform pressure P on the piston and the walls of the cylinder. If the cross-sectional area of the piston is A, the force exerted by the gas on the piston is

$$F = P \times A$$

When the gas expands quasistatically as the piston goes up a distance dy as in Fig.1.2(b), the work done by the gas on the piston is

$$\delta W = F \times dy = PAdy$$

But $Ady = dV$, the increase in the volume of the gas. Therefore,

$$\delta W = P\,dV \tag{2}$$

The above equation gives the work exclusively in terms of the thermodynamic variables of the system. The nature of the outside force and other characteristics of the surroundings are not reflected in this relation. The work $\delta W = P\,dV$ is often called thermodynamic work.

The convention is that work done by the system is regarded as positive and the work done on the system is negative. Thus, the work done by the gas expanding against a piston is positive and the work done by a piston compressing a gas is negative.

The total work done by the gas as its volume changes from V_1 to V_2 can be found by integrating equation (2). Thus,

$$_{1}^{2}W = \int_1^2 \delta W = \int_1^2 PdV \tag{3}$$

The above integration can be performed only if we know the relationship between P and V during the process. In general, the pressure is not constant but depends on the volume and temperature. If the pressure and volume are known at each step of the process, the work done can be obtained graphically from the PV diagram as in Fig.1.3.

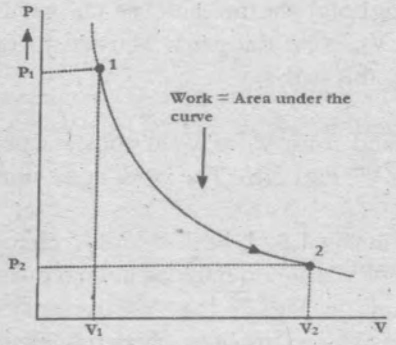


Fig. 1.3. A gas expands from an initial state 1 to final state 2. The work done by the gas equals the area under the PV curve

The work done during the process is given by the area V_1-1-2-V_2-V_1 under the curve 1-2 in Fig.1.3. Therefore, the total work done during the expansion of the gas from the initial state to the final state is the area under the curve in a PV diagram.

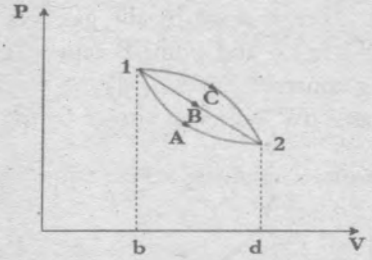


Fig 1.4. Three different paths connecting the same initial and final states. The work done by the gas is maximum for the path marked C which encloses the largest area.

As illustrated in Fig1.4. it is possible to go from the initial state 1 to the final state 2 along any path like A, B or C. The area under each curve symbolizes the work for each process. The amount of work done is a function of the end states of the process and on the path that is followed. To illustrate this important point we consider three different paths (Fig 1.5) connecting initial and final states.

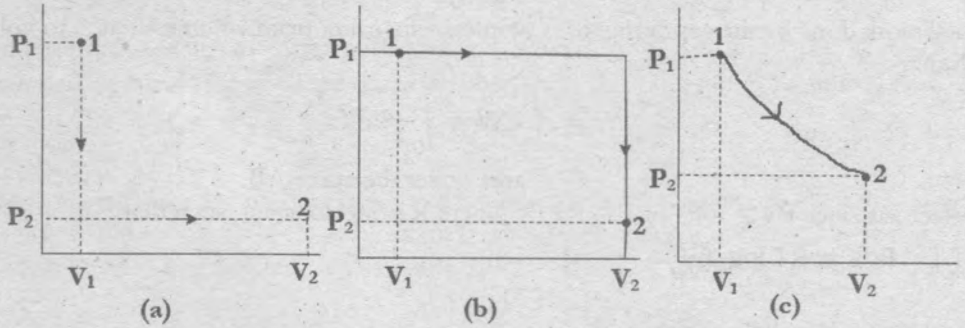


Fig. 1.5 The work done by an ideal gas as it is taken from an initial state 1 to final state 2 depends on the path between these states 1 and 2.

In the process represented in Fig1.5(a) the pressure of the gas is first reduced from P_1 to P_2 by cooling it at a constant volume V_1. Next the gas is allowed to expand from V_1 to V_2 at constant pressure P_2. The work done along this path is

$W_A = P_2(V_2 - V_1)$

The gas is first permitted to expand from V_1 to V_2 at constant pressure P_1 and then the pressure is reduced to P_2 at constant volume V_2, Fig1.5(b). The work done along this path is

$W_B = P_1(V_2 - V_1)$

Finally in the process described in Fig 1.5(c), both P and V change continuously. To complete the work in this case, the shape of the PV diagram must be known unambiguously.

It is clear from the PV diagrams in Fig1.5, that W_A is smaller than W_B and W_C has a value intermediate between W_A and W_B. This example amply demonstrates that the work done by a system depends on how the system goes from the initial to the final state. For this reason, thermodynamic work is not a point function but is a path function. dW cannot be therefore treated as exact differential in the mathematical sense.

1.8.1 Work done in Various Processes

(i) Isothermal Process

When a gas expands isothermally, work is done by the gas. Let P_1, V_1 be the initial pressure and volume represented by point A in Fig.1.6 and point B represents P_2, V_2 as the final pressure and volume, the temperature remaining constant. Suppose dV be the small increase in the volume of gas at pressure P. Work done by the gas is $dW = PdV$ as shown by the shaded area in Fig. 1.6

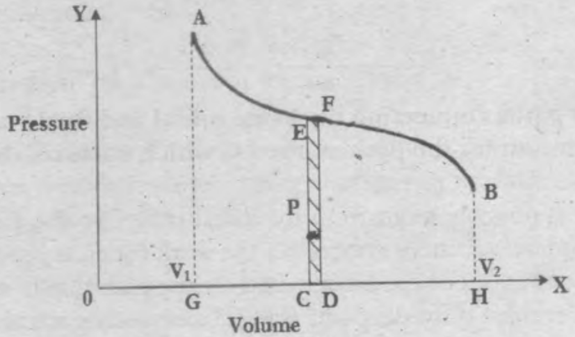


Fig. 1.6

Thus, total work done by the gas during the complete expansion from volume V_1 at A to volume V_2 at B such that

$$W = \int dW = \int_{V_1}^{V_2} PdV$$

= area under the curve AB.

For a perfect gas since $PV = RT$ or $P = RT/V$ where R is the constant, such that

$$W = RT\int_{V_1}^{V_2} PdV = RT\log_e\left(\frac{V_2}{V_1}\right)$$

Or

$$W = 2.3026\ RT\log_{10}\left(\frac{V_2}{V_1}\right) \tag{4}$$

As the temperature is constant, thus

$P_1V_1 = P_2V_2$ or $\frac{V_2}{V_1} = \frac{P_1}{P_2}$

From (4), we have

$$W = 2.3026\ RT\log_{10}(\frac{P_1}{P_2}) \quad (5)$$

As the temperature remains constant, the internal energy of the gas undergoes no change. The whole of the heat applied to the gas is thus converted into external work done by the gas.

(ii)**Adiabatic Process**

Let us consider an adiabatic expansion of a 1 gm mole of a perfect gas having P_1, V_1 as the initial pressure and volume and P_2, V_2 as the final pressure and volume.

Work done by the gas is: $W = \int dW = \int_{V_1}^{V_2} PdV$

During an adiabatic process,

$PV^{\Upsilon} = k$ a constant

Or $P = \frac{k}{V^{\Upsilon}}$

On substituting the value of P above we get,

$$W = \int_{V_1}^{V_2} \frac{k}{V^{\Upsilon}} dV = k\int_{V_1}^{V_2} V^{-\Upsilon} dV = k\,[\frac{V^{1-\Upsilon}}{1-\Upsilon}]_{V_1}^{V_2}$$

$$= \frac{k}{1-\Upsilon}\left[V_2^{1-\Upsilon} - V_1^{1-\Upsilon}\right]$$

Or

$$W = \frac{k}{\Upsilon-1}\left[\frac{1}{V_1^{\Upsilon-1}} - \frac{1}{V_2^{\Upsilon-1}}\right] \quad (6)$$

If V_2 is smaller than V_1, work will be done on the gas and the above expression (6) will then bear a negative sign.

Also, $P_1V_1^{\Upsilon} = k = P_2V_2^{\Upsilon}$, hence from (6) above

$$W = \frac{1}{\Upsilon-1}\left[\frac{P_1V_1^{\Upsilon}}{V_1^{\Upsilon-1}} - \frac{P_2V_2^{\Upsilon}}{V_2^{\Upsilon-1}}\right]$$

Or $$W = \frac{1}{\Upsilon-1}[P_1V_1 - P_2V_2] \quad (7)$$

Again, since $P_1V_1^{\Upsilon} = k = P_2V_2^{\Upsilon}$, we have $V_1 = [\frac{k}{P_1}]^{1/\Upsilon}$ and $V_2 = [\frac{k}{P_2}]^{1/\Upsilon}$

Thus, $$W = \frac{1}{\Upsilon-1}[P_1[\frac{k}{P_1}]^{\frac{1}{\Upsilon}} - P_2[\frac{k}{P_2}]^{1/\Upsilon}]$$

Or $$W = \frac{k^{1/\Upsilon}}{\Upsilon-1}[P_1^{(\Upsilon-1)/\Upsilon} - P_2^{(\Upsilon-1)/\Upsilon}] \quad (8)$$

If temperature of the gas changes from T_1 to T_2 we have $P_1V_1 = RT_1$ and $P_2V_2 = RT_2$ where R is the gas constant of gm-molecule of the gas. On substituting the values of P_1V_1 and P_2V_2 in (7) above we get,

$$W = (\frac{R}{\Upsilon-1})(T_1 - T_2) \quad (9)$$

Since in an adiabatic process, no heat is allowed to enter or leave the system, the external work W is done by the gas at the expense of its own internal energy, such that

Work done by the gas during an adiabatic process = decrease in internal energy of the gas

Thus, the work done by the gas during an adiabatic expansion from volume V_1 at pressure P_1 to volume V_2 at pressure P_2 is given by the area under the P-V curve for the gas i.e.

$$W = \int_{V_1}^{V_2} PdV = \frac{1}{\Upsilon-1}[P_1V_1 - P_2V_2] \doteq \text{area under the curve.}$$

1.9 Heat in Thermodynamics

A process of transfer of energy without work being done is called heat exchange.

Definition: Heat is a form of energy that is moved from a system at higher temperature to another system at a lower temperature.

So, heat is energy in transit. Heat transferred to a system is considered positive and from a system is negative.

Like work, heat is a path function. The heat transferred to or from a system depends on the process. We illustrate this point with the help of following examples:

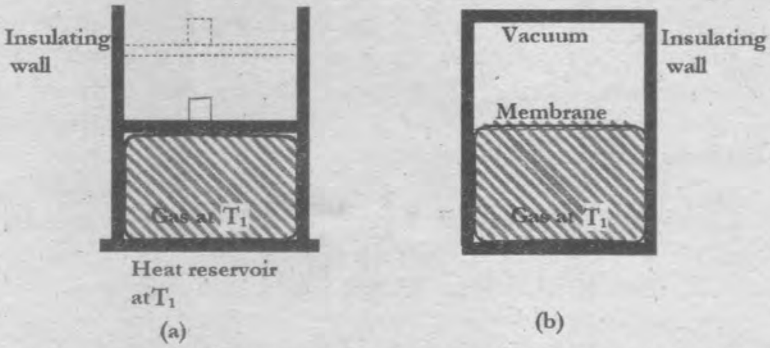


Fig.1.7(a) A gas at temperature T_1 expands slowly by absorbing heat from the reservoir at the same temperature (b)A gas expands rapidly into the evacuated region due to the rupture of the membrane.

Let us consider the case of a gas of volume V_1 which is in thermal contact with a heat reservoir Fig 1.7 (a) If the pressure of the gas is infinitesimally greater than atmospheric pressure the gas will expand and pushes up the piston. Ultimately the gas expands to a final volume V_2. The heat required to maintain the gas at a constant temperature T, during the process of expansion, is supplied from the reservoir to the gas. Work is done by the gas in this case.

In the second case, illustrated in Fig.1.7(b), a gas of the same volume V_1 is thermally insulated. It can neither receive nor give away heat. When the membrane is broken, the gas freely and rapidly expands into the vacuum until it occupies a volume V_2. No work is done in this case. Further, no heat is transferred through the walls. The initial and final states of both processes are identical but the path followed are different. We conclude that heat, like work, depends on how the system goes from its initial state to final state. Therefore, it is a path function. Like δW, δQ is an inexact differential.

1.10 Comparison of Heat and Work

In thermodynamics, work includes all forms of energy except heat. When a battery is charged, electrical work is performed and stored in chemical form. When a piece of steel is magnetized, magnetic work is performed.

Any kind of energy in the course of transformations may pass through many forms but invariably ends in the form of heat energy. In the process of mechanical motion, the kinetic energy of a body decreases owing to the action of frictional forces and gets transformed into heat. Similarly, the energy of an electric current and of chemical reactions transform into heat.

Both work and heat are transient phenomenon. Systems never contain work or heat. Heat and work exist only in a process of heat energy transfer and their numerical value depends on the kind of the

process. Either work or heat or both are transferred when a system undergoes a change of state. In real conditions, both ways of transferring energy to a system accompany each other. For instance, when a metal rod is heated, heat exchange takes place and at the same time thermal expansion of rod occurs. The latter implies that the work of expansion is done.
Heat and work are not equivalent forms of energy transfer from a qualitative viewpoint. The energy of ordered motion is transferred in the form of work. The energy of chaotic motion of particles constituting the body is increased when energy is transferred to a body in the form of heat.
Work can be fully converted into heat but heat cannot be entirely converted into work. As both heat and work depend on the path, neither quantity is individually conserved during a thermodynamic process.

1.11 Internal Energy

The total energy E_T of a system consists of (i)the kinetic energy of its macroscopic motion as a whole(ii)the potential energy due to the presence of external fields and (iii)internal energy, U. Thus,

$$E_T = K.E + P.E + U \tag{10}$$

The internal energy U of the system depends on the nature of the motion and interaction of the particles in the system. It consists of (a)the kinetic energy of random thermal motion of molecules (b)the potential energy of molecules due to the intermolecular interactions, (c)the kinetic and potential energies of atoms and electrons and (d) the nuclear energy. In thermodynamics, we do not concern ourselves with the form of internal energy.
Experiments have shown that the internal energy is determined by the thermodynamic state of the system and does not depend on how the system acquires the given state. Consequently, the internal energy is not related to the process of a change in the state of the temperature. All processes leading to particular pressure and temperature leave the gas with the same internal energy.
The internal energy is not of practical interest. The change in internal energy ΔU, when a system changes from one state to another is of actual interest. It is generally assumed that the internal energy of a system is zero at 0 K.

1.12 Law of Conservation of Energy

In a closed mechanical system, the sum of kinetic and potential energy is constant. The total energy can neither be created nor destroyed. It means that the energy is conserved. This is the law of conservation of mechanical energy. This law is applicable only in situation when there is no transformation of mechanical energy into heat energy. In real systems, the mechanical motion in general is accompanied with heating. When the engine of a running car is switched off, it gradually slows down and ultimately stops. Apparently, its kinetic energy has disappeared. In fact, its kinetic energy has transformed into heat energy by the frictional forces. Both, the tyre of the car and the ground are heated up. As a result, the random motion of particles constituting the interacting tyres and road acquires more velocity. To sum up, the mechanical energy (K. E of car) is transformed into the internal energy of the interacting bodies. If we take into account the internal energy, the law of conservation of energy can be extended to include thermodynamic systems also. The first law of thermodynamics generalizes the law of consideration of mechanical energy.

1.13 First Law of Thermodynamics

Thermodynamics involves the three general forms of energy i.e., heat, work and internal energy.

- Heat is energy transferred by virtue of temperature difference.

- Thermodynamic work is energy transferred between system and surroundings
- Internal energy is energy stored within a system

If the state of a system changes as a result of supplying a quantity of heat Q to it and as a consequence the system does the work W, then the law of conservation of energy states that the quantity of heat supplied to the system will be equal to the sum of the work performed by the system and the change in the internal energy of the system. That is,
Net heat transfer = Work + change in internal energy
Or mathematically, $Q = W + \Delta U$ (11)
Above equation (11) is the first law of thermodynamics and is a consequence of the conservation of energy. While the quantities Q and W are path dependent, the internal energy, does not depend on the path of the process.
Suppose a thermodynamic system undergoes a change from an initial state 1 to a final state 2 in which Q units of heat are absorbed (or removed) and W is the work done by (or on) the system. Expressing both Q and W in the same units (either thermal or mechanical) the difference (Q – W) can be calculated. If now we carry out this calculation for different paths between the same states 1 and 2, the quantity (Q – W) will be the same for all paths connecting the states 1 and 2. It follows that the internal energy change of a system is independent of the path. If U_1 is the internal energy in state 1 and U_2 is the internal energy in state 2, then the equation (11) can be written as
$U_2 - U_1 = \Delta U = Q - W$ (12)
When a thermodynamic process proceeds smoothly, it can be treated as a continuous sequence of small changes. Mathematically, we write equation (11) in the differential form as:
$dQ = dW + dU$ (13)
It is important to note that expressing heat and work as dQ and dW does not imply the existence of properties Q and W that measure the heat and work content of a system. dQ and dW denote small amounts of heat and work but they are not true differentials:
There is a series limitation on the first law of thermodynamics. The law tells us whether energy considerations permit a particular process to take a system from one equilibrium state to another. It does not tell us whether this process will actually occur or not. A certain process might be entirely consistent with the principle of energy conservation but still it will not take place.

1.14 Applications of the First Law
Let us study some special processes and the application of the first law to them
(i)Isolated system: An isolated system does not interact with its surroundings. Therefore, there is no heat flow and the work done is zero. That is $Q = 0$ and $W = 0$. It follows from equation (12) that
$\Delta U = 0$
$U_2 - U_1 = 0$
or $U_2 = U_1$ isolated system (14)
equation (14) means that the internal energy of an isolated system remains constant.
(ii)Cyclic process: In a cyclic process, the initial and final states of the system are the same. Thus,
$U_2 = U_1$
$\Delta U = 0$
It follows from equation (11) that

$$Q - W = 0$$

$Q = W$ cyclic process (15)

Above equation (15) means that the net work done by the system over a cycle equals the net heat absorbed over the cycle.

(iii)Adiabatic process: A process in which no heat is absorbed or ejected by the system is called an adiabatic process. Thus, $\Delta Q = 0$

$\Delta U = -W$ Adiabatic process (16)

Thus, the change in the internal energy of the system is equal in magnitude to the work done by the system. Heat flow into the system from the surroundings may be prevented in two ways- (i)by surrounding the system with a thick layer of insulating material or (ii)by performing the process quickly. The flow of heat requires finite time. So, any process performed quickly enough will be adiabatic.

(iv)Isochoric process: When a substance undergoes a process in which the volume remains unchanged, the process is called isochoric. If the volume of a system remains constant, it can do no work. Thus, $W = 0$ and the first law gives

$\Delta U = Q$ Isochoric process (17)

In this case, the heat that entered the system is stored as internal energy.

(v)Isothermal process: It is process happening at constant temperature and the quantities Q, W and ΔU are all non-zero. The first law does not have any special form for an isothermal process.

(vi)Isobaric process: A process taking place at constant pressure is called an isobaric process. As in the case of an isothermal process, Q, W and ΔU are all non-zero. The work done by a system that expands or contracts isobarically has a simple form. As the pressure is constant,

$$W = \int_1^2 P\,dV = P(V_2 - V_1) \quad \text{isobaric process} \tag{18}$$

(vii)Isothermal expansion of an ideal gas: Let an ideal gas on expansion in quasistatic process at constant temperature by placing the gas in good thermal contact with a heat reservoir. The ideal gas equation $PV = nRT$ for each point on the path

$$W = \int_{V_1}^{V_2} P\,dV = \int_{V_1}^{V_2} \frac{nRT}{V}\,dV$$

As T is constant in the process, we can write

$$W = nRT \int_{V_2}^{V_1} \frac{dV}{V}$$

$$W = nRT\ln\left(\frac{V_2}{V_1}\right) \tag{19}$$

(viii)Adiabatic expansion of an ideal gas:

Let us find the relation between P and V for an adiabatic process carried out on an ideal gas. According to equation (16), $\Delta U = -W$ in an adiabatic process

For an ideal gas $\Delta U = nC_v dT$ (at constant volume)

The work done during the process is given by $W = P\,dV$

$$P\,dV = -nC_v dT \tag{20}$$

We can write the equation of the state of the gas in differential form as

$$d(PV) = d(nRT)$$

so

$$P\,dV + V\,dP = nR\,dT \tag{21}$$

Using equation (20) into equation (21), we get

$$V\,dP = nC_v dT + nRdT\,(C_v + R)$$

But $C_v + R = C_P$

$V\,dP = nC_P dT$ (22)

Taking the ratio between equation (22) and equation (21), we get

$$\frac{V\,dP}{P\,dV} = -\frac{C_p}{C_v} = -\gamma$$

$$\frac{dP}{P} = -\gamma\frac{dV}{V}$$

Integrating on both the sides of the above equations, we get

$$\int_{P_1}^{P_2}\frac{dP}{P} = -\gamma\int_{V_1}^{V_2}\frac{dV}{V}$$

Or $\ln\frac{P_2}{P_1} = -\gamma\ln\frac{V_2}{V_1}$

Thus, $P_1V_1^\gamma = P_2V_2^\gamma$ (23)

Or $PV^\gamma = \text{constant}$ (24)

1.15 Heat Engine

A thermodynamic device that converts heat supplied to it into mechanical work is called heat engine. The automobile engines and steam turbines are examples of heat engines. Irrespective of design and features, for theoretical purposes, any heat engine can be conveniently represented by a diagram as shown in Fig.1.8

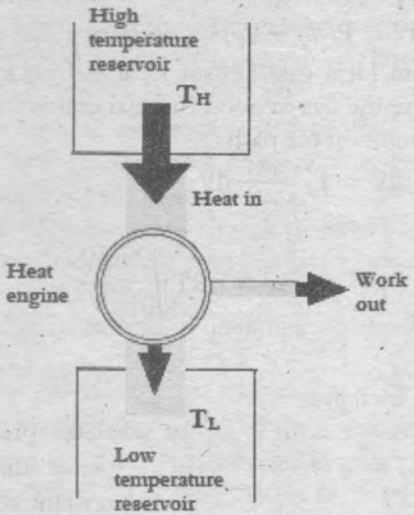


Fig. 1.8 Schematic representation of heat engine. The engine absorbs heat Q_1 from the high temperature reservoir, expels heat Q_2 from the high temperature reservoir and does work W.

Any heat engine consists of three main parts:

1.**Source:** There must be a source of heat or a hot body or a hot reservoir with infinite thermal capacity maintained at high temperature so that any amount of heat given or taken away does not affect its temperature.

2.**Sink:** There must be a condenser or sink or cold reservoir with infinite thermal capacity maintained at a lower constant temperature so that whenever any amount of heat is given or taken away does not affect its temperature.

3.**Working substance:** There must be a working medium or a substance through which heat can be absorbed from the source and rejected to sink. e.g. in case of steam engine, its steam and air in case of internal combustion engines of scooters, motor cars, aeroplanes etc.
The difference between the quantities of heat absorbed by the working substance from the source and rejected to the sink is converted into work (barring the heat lost in overcoming friction or by conduction etc). This means that the working substance falls in temperature during the process.

1.15.1 Thermal Efficiency

Practically no heat engine can convert whole of the heat supplied into work, a part of heat is usually rejected as unused heat. If the heat energy supplied is Q_1 and the heat energy rejected is Q_2 the amount of heat utilized in doing work is $Q_1 - Q_2$. Efficiency is thus defined as the ratio between its output of work and the input of heat and is denoted by η. It is a measure of how economical an engine is. We express cost efficiency in terms of fuel economy.

If the heat energy supplied is Q_1 and the heat energy rejected is Q_2 the amount of heat utilized in doing work and W is the work done by the engine, we have

Efficiency of the engine, $\eta = \frac{W}{Q_1}$

Since after every cycle, the working substance regains its original state the internal energy is not affected. Under ideal conditions there will be no dissipation of energy due to friction or conduction etc. We have,

$$W = Q_1 - Q_2 \tag{25}$$

$$\eta = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{Q_2}{Q_1}$$

It is expressed as percentage.
For an ideal engine,

$$\eta = 1 - \frac{Q_2}{Q_1} \tag{26}$$

In practice, the useful work delivered by an engine is less than the work W owing to friction losses. Therefore, the overall efficiency is less than the thermal efficiency. Thus, for real engine

$$\eta \leq 1 - \frac{Q_2}{Q_1}$$

Where the less than sign refers to real engines and the equal sign to an ideal engine in which there are no losses.

1.16 The Carnot Cycle and Carnot Engine

In 1824 French Engineer, Sadi Carnot conceptualized an ideal engine and demonstrated that a heat engine operating in a reversible cycle between the two heat reservoirs would be the most efficient engine possible. This ideal engine is called Carnot engine.
The net work done by the working substance taken through the Carnot cycle is the largest possible for given amount of heat supplied to the substance.
The Carnot cycle is an idealization of the cycle of a real heat engine. It is assumed that there are no losses of energy by heat exchange with the environment and that there is no friction and other imperfections of actual engines. It must absorb all its heat at a constant high temperature and can reject whatever heat it has to at a constant low temperature and should work in a cycle of operations each of which is perfectly reversible.
He assumed the following things:

1.A cylinder of perfectly non-conducting walls with a perfectly conducting bottom fitted with a perfectly non-conducting and frictionless piston. The working substance in the piston can be assumed to be air which behaves like a perfect gas.

2.A hot reservoir or source of heat with infinite thermal capacity maintained at a high and constant temperature T_1.

3. A cold reservoir or sink with infinite thermal capacity maintained at a lower and constant temperature T_2.

4.A perfectly non-conducting stand on which when desired the cylinder can be moved to it without any friction.

The working substance has to undergo following quasi-static operations in order to obtain a continuous supply of work:

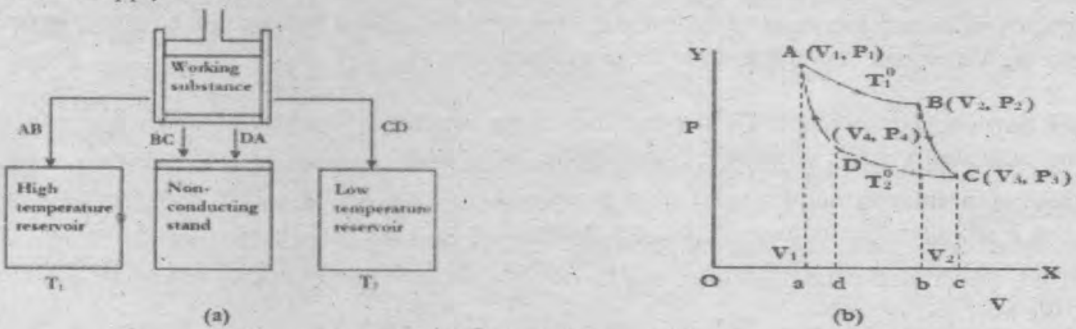


Fig. 1.9. The Carnot cycle

1.The cycle begins at the equilibrium state A Fig.1.9(b) with the cylinder in contact with heat source at a temperature T_2 Fig. 1.9(a). The working substance (the ideal gas) undergoes a slow quasistatic isothermal expansion to state B. As the gas expands it does external work against the pressure on the piston and its temperature tends to decrease. The fall in temperature is immediately made-up by-passing heat into the cylinder through its perfectly conducting base which is in contact with the source. Let the total amount of heat absorbed by the gas from the source during the process be Q_1. This heat energy input is converted directly to the work W_1 done in the moving piston. This operation is represented by the isothermal AB on the indicator diagram, shown in Fig 1.9(b) where A represents the initial volume V_1 and pressure P_1 of the gas and B, its final volume V_2 and pressure P_2 after absorbing Q_1 units of heat from the source and performance of W_1 units of external work.

From the first law of thermodynamics, the quantity of heat absorbed Q_1 must be equal to the external work W_1 done by the gas during isothermal expansion along AB from volume V_1, pressure P_1 to volume V_2, pressure P_2 at temperature T_2. This is represented by ABbaA. So,

$$Q_1 = W_1 \int_{V_1}^{V_2} PdV$$

$$= RT_1 \log_e \frac{V_2}{V_1} = \text{area ABbaA} \qquad (27)$$

2.With the system at B, the cylinder is removed from contact of source i.e., from the high temperature reservoir and placed on the non-conducting stand. This makes the gas thermally isolated from the surroundings. The gas slowly expands adiabatically along the path B to C to the state C. During this part of the cycle B to C no heat leaves or enters the system. However, work W_2 is done at the expense of reducing the internal energy. Consequently, the temperature falls from T_1 to T_2.

This operation is represented by adiabatic BC where C represents the volume V_3 and pressure P_3 of gas at temperature T_2. So that W_2 will be the work done by the gas it is represented by the area BCcbB on the indicator diagram. We thus have,

$$W_2 = \int_{V_1}^{V_2} PdV = \frac{R(T_1 - T_2)}{\gamma - 1} = \text{area BCcbB} \tag{28}$$

3.Next the cylinder is placed in contact with the low temperature reservoir i.e., sink at a temperature T_2. The gas is now slowly compressed (by an external agency here piston) so the work is done on the gas. During this part of cycle C to D, the temperature and hence the internal energy of the working substance are constant. The gas expels heat Q_1 to the reservoir and the work done on the gas by the external agent is W_3. This is shown here by the isothermal CD where D represents the volume V_4 and pressure P_4 of the gas when it has rejected Q_2 units of heat to the sink and W_3 units of work has been done upon it represented as by the area CcdDC. Then we have,

$$Q_2 = W_3 = \int_{V_3}^{V_4} PdV = RT_2 \log_e \frac{V_4}{V_3} = -RT_2 \log_e \frac{V_3}{V_4} \tag{29}$$

$$= \text{area CcdDC}$$

4.In the final part D to A, the cylinder is removed from the contact with the low temperature reservoir or sink and again placed on the non-conducting stand. The gas is compressed adiabatically and brought to the initial state A. The adiabatic compression is also the result of work W_4 done on the gas by the external energy. Consequently, to the adiabatic compression the system temperature increases from T_1 to T_2.

This step completes the Carnot cycle and returns the system to the initial condition.

The operation is represented by adiabatic DA and the work done on the gas is represented as DdaAD.

$$W_4 = \int_{V_4}^{V_1} PdV = -\frac{R(T_1 - T_2)}{\gamma - 1} = \text{area DdaAD} \tag{30}$$

Thus, it can be concluded that the Carnot cycle consist of two isothermal and two adiabatic strokes one each of compression and expansion, the two alternating with each other.

Work done by the engine per cycle:

The net work done in this reversible process is equal to the area enclosed by the path ABCDA of the PV diagram Fig 1.9(b). In a reversible cycle, the change in internal energy is zero. It follows from first law that the net work done in one cycle equals the net heat transferred into the system. For a system that undergoes a Carnot cycle, no heat is supplied to or rejected by the system during the adiabatic paths BC and DA as in Fig 1.9(b). An amount of heat Q_1 is supplied to the system during the isothermal expansion AB and an amount Q_2 is rejected during the isothermal compression CD. Thus, the first law can be written as:

$W = Q_1 - Q_2$ all heat engines

Above equation is applicable for a complete cycle

Of any heat engine the thermal efficiency is given by:

$$\eta = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{Q_2}{Q_1}$$

1.16.1 Efficiency

Let us now calculate Efficiency (η) in terms of the temperatures T_1 and T_2 of the source and sink respectively and in terms of the expansion ratio (ρ) of the gas.

I

n terms of the temperatures T_1 and T_2 of the source and sink

$$\frac{Q_1}{Q_2} = \frac{W_1}{W_2} = \frac{RT_1 \log_e \frac{V_2}{V_1}}{RT_2 \log_e \frac{V_3}{V_4}} = \frac{T_1 \log_e \frac{V_2}{V_1}}{T_2 \log_e \frac{V_3}{V_4}} \quad (31)$$

Since points B and C lie on the same adiabatic BC, we have

$$T_1 V_2^{\gamma-1} = T_2 V_3^{\gamma-1} \text{ or}$$

$$\frac{T_1}{T_2} = \left[\frac{V_3}{V_2}\right]^{\gamma-1} \quad (32)$$

Where γ is the ratio between Cp and Cv for the gas.

Similarly, points D and A lie on the same adiabatic DA, we have

$$T_1 V_1^{\gamma-1} = T_2 V_4^{\gamma-1} \text{ or}$$

$$\frac{T_1}{T_2} = \left[\frac{V_4}{V_1}\right]^{\gamma-1} \quad (33)$$

From equation (26) and equation (27) :

$$\frac{V_3}{V_2} = \frac{V_4}{V_1} \quad \text{or} \quad \frac{V_2}{V_1} = \frac{V_3}{V_4}$$

Thus,

$$\log_e \frac{V_2}{V_1} = \log_e \frac{V_3}{V_4} \quad (34)$$

Substituting in equation (31) we have,

$$\frac{Q_1}{Q_2} = \frac{T_1}{T_2} \quad (35)$$

Thus, efficiency,

$$\eta = 1 - \frac{Q_2}{Q_1} = 1 - \frac{T_2}{T_1}$$

Efficiency always depends on the temperatures T_1 and T_2 of the source and sink respectively and that the greater the value of $(T_1 - T_2)$, greater the difference between the temperatures of the source and sink, higher will be the efficiency. Since, $(T_1 - T_2)$, must always be less than T_1, it is clear that the efficiency must be less than 1 or 100%

In terms of expansion ratio:

We have seen here that

$$\frac{V_3}{V_2} = \frac{V_4}{V_1} \quad (36)$$

Each one of the ratios is the adiabatic expansion ratio of the gas, and so is equal to ρ

From equation (32) and (33),

$$\rho = \frac{V_3}{V_2} = \frac{V_4}{V_1} = \left[\frac{T_1}{T_2}\right]^{1/\gamma-1} \text{ whence, } \frac{T_2}{T_1} = \left[\frac{1}{\rho}\right]^{\gamma-1} \quad (37)$$

Thus efficiency,

$$\eta = 1 - \left[\frac{1}{\rho}\right]^{\gamma-1} \quad (38)$$

1.16.2 Carnot's Theorem

This theorem discloses a fundamental restriction on the conversion of heat into work. It is the most efficient heat engine operating between the two heat reservoirs.

Carnot's Theorem Statement: No heat engine can be more efficient than a reversible Carnot engine operating between the same limits of temperature i.e., between the same source and sink and all reversible engines operating between the same limits of temperature have the same efficiency.

1.17 Heat Pump

A device that transfers heat from a sink to a high temperature reservoir i.e., source is known as heat pump. This is the reverse function of a heat engine. It requires work input, because the heat transfer from low temperature reservoir to high temperature reservoir will not take place spontaneously.
The household refrigerators and air conditioners are examples of the most common heat pumps. In an ordinary household refrigerator, the working substance is a liquid that circulates within the system. It takes heat from the low temperature reservoir, which is the cold chamber in which food article are stored and transfers it to outside air in the room where the unit is kept. The work is supplied by a compressor that uses electrical energy. The working substance is a fluid that readily undergoes a liquid to gas phase change at the operating temperature. The most common refrigerants are Freon-12 (CCl_2F_2) and ammonia.

1.18 Second Law of Thermodynamics

The first law of thermodynamics is called the law of equivalence and can only tell us about the inter-convertibility of heat and mechanical work. It tells us nothing about the conditions in which the conversion can occur neither it tells us the limit within which it is possible. Different workers have defined it in different ways. Here we can discuss two classical statements of the second law of thermodynamics: Kelvin-Planck statement and Clausius statement.

According to Kelvin-Planck: It is impossible to construct a heat engine that will operate in a cycle and which will receive a given amount of heat from a high temperature reservoir and does an equal amount of work. The only alternative is that some heat must be transferred from the working fluid to a low temperature reservoir. Therefore, work can be done by the transfer of heat only if there are two temperature levels involved.

According to Rudolf Clausius: It is impossible to construct a device that operates in a cycle and produce no effect other than the transfer of heat from a cold body to the hot body. Thus, it is impossible to construct a refrigerator that operates without an input of work.
The first law of thermodynamics is concerned with conservation of energy. As long as the energy is conserved in a process, the first law is satisfied. The second law underlines the fact that the processes proceed in a certain direction but not in the opposite direction. For example: a cup of tea cools by heat transfer to the surroundings but heat will not flow from the cooler surroundings to hot the tea cup. Thus, real processes progress only in one direction.
A cycle will occur only when both the first and second law of thermodynamics are satisfied.
The first law of thermodynamics is a common declaration of the conservation of energy. It makes no distinction between the different forms of energy. The second law of thermodynamics asserts that thermal energy is different from all other forms of energy. Various forms of energy can be converted into thermal energy spontaneously and completely; whereas the reverse transformation is never complete. The impossibility of converting heat completely into mechanical energy forms the basis of Kelvin-Planck statement of second law.
The basis of second law lies in the difference between the nature of mechanical energy and the nature of internal energy. Mechanical energy is the energy of ordered motion of a body. Internal energy is the energy of random motion of molecules within it as a whole in the direction of velocity of the body. The energy associated with this ordered motion of molecules is the internal energy. When a moving body comes to rest due to friction, the ordered portion of the kinetic energy becomes converted into energy of random molecular motion. It is impossible to reconvert the energy of

random motion completely to the energy of ordered motion, since we cannot control the motions of individual molecules. We can convert only a portion of it. That is what a heat engine does.

1.19 Entropy

The differential form of the first law of thermodynamics is written as

$$dQ = dU + dW \tag{39}$$

The work dW done depends on the path of process and therefore it is not a function of the state. The same is the case with dQ, the quantity of heat supplied or taken away. The work dW can be expressed in terms of thermodynamic variables and their changes. For instance, we have expressed dW in equation (39) as

$$dW = PdV \tag{40}$$

It is found that dQ can also be expressed in a similar fashion, in case of reversible processes. We write

$$dQ = T\,dS \tag{41}$$

where dS is called the change in entropy and T is the temperature. Now, we define the change in entropy as

$$dS = dQ/T \qquad \text{reversible process} \tag{42}$$

The change in entropy dS in the course of an infinitesimal change is equal to the quantity of heat dQ divided by the absolute temperature T, where dQ is the heat absorbed or rejected when the change is carried out in a reversible manner.

The total entropy change in a reversible process may be obtained by integrating equation (42) thus,

$$\Delta S = S_2 - S_1 = \int_1^2 \frac{dQ}{T} \tag{43}$$

Where S_1 and S_2 are the entropies of the initial and final states of the system.

The importance of the entropy S is that it is a function of state like the internal energy U. Both these parameters depend only on the initial and final states of the system and not on the path of the process that takes the system from the initial state to final state. Equation (43) assumes a simpler form when the process is an isothermal process. As T is constant in isothermal process equation (43) may be written as

$$\Delta S = \int_1^2 \frac{dQ}{T} = \frac{1}{T}\int_1^2 dQ = \frac{Q}{T} \tag{44}$$

Thus, $\Delta S = \frac{Q}{T}$ isothermal process

The units of entropy and entropy change are J/K

In practice, the value of entropy S is not of much interest. We have to know the change in the entropy when the system changes from one state to another.

1.19.1 Entropy, Disorder and Second Law

When processes occur, in general they are irreversible and the degree of disorder increases as a result of these processes. As an example, let us take the case of isothermal expansion of an isothermal gas. As the gas absorbs heat, it slowly expands. At the end of the process the gas occupies a greater volume than at the beginning. The gas molecules are more disordered now. The gas will not, by its own accord, give up its thermal energy and segregate itself to confine to the initial volume. We thus, observe that the flow of heat takes place in the direction that increases the amount of disorder. The same type of order to disorder change occurs when free expansion of gas occurs, when one gas diffuses into another, and in similar other spontaneous processes.

Rudolf Clausius the German physicist, introduced the quantity entropy which is regarded as a measure of disorder in a system. An increase in disorder is equivalent to an increase in entropy. Irreversible processes are processes for which entropy increases.

These considerations led Clausius to reformulate the second law of thermodynamics in terms of entropy. According to it, the entropy of an isolated system always tends to increase. Mathematically, it is expressed as

$$\Delta S_{\text{isolated system}} \geq 0$$

1.19.2 Points of Discussion

(i)The net change in entropy in any reversible cycle is zero

Let us take the case of Carnot cycle as an example. There is no change in entropy of the working substance during two adiabatic paths. Either during adiabatic expansion or compression $Q = 0$. Therefore, $S = 0$. However, there is an increase in entropy during isothermal expansion as heat Q_1 is added at a constant temperature T_1. The consequent increase in entropy $\Delta S_1 = \frac{Q_1}{T_1}$. There is a decrease in entropy during isothermal compression in which heat Q_2 is rejected at a temperature T_2. Thus, $\Delta S_2 = -\frac{Q_2}{T_2}$

The net change in entropy is given by

$$\Delta S = \Delta S_1 + \Delta S_2 = \frac{Q_1}{T_1} - \frac{Q_2}{T_2}$$

But $\frac{Q_1}{T_1} = \frac{Q_2}{T_2}$

Thus, ?? $\Delta S = 0$ reversible cycle

(ii)Entropy increases in all reversible processes: It is proved that there is net increase in the entropy during irreversible processes. Since all real processes taking place in the universe are irreversible, there is a continuous increase in its entropy. For this reason, entropy is not conserved. In this respect entropy differs from energy.

We can illustrate this by taking example. Suppose a small quantity of heat dQ is radiated away from a hot body A, at a temperature T_H to a cold body B at a temperature T_C. Let dQ be so small that T_H and T_C are not altered appreciably due to the exchange of heat. However, the entropy of A decreases by $-dQ/T_H$ whereas that of B increases by dQ/T_C in this process.

$$\Delta S = \frac{dQ}{T_C} - \frac{dQ}{T_H}$$

As $T_H > T_C$, $\Delta S > 0$

(iii)Entropy indicates the direction in which processes proceed in nature: all natural processes are irreversible. They proceed in the direction of increasing entropy.

(iv)Entropy represents the unavailability of energy: In the thermodynamic sense, entropy is a measure of the capability to do work or transfer heat. A system at a higher temperature will tend to do work on and/or transfer heat to its lower temperature surroundings. In the process, the entropy of the system increases and greater the entropy the less available is the energy.

Let us consider the example of Carnot engine. The efficiency of Carnot engine is given by

$$\eta = 1 - \frac{T_2}{T_1}$$

As Q_H is the heat input, heat converted into work $= Q_H(1 - \frac{T_2}{T_1})$

Thus, heat unavailable for work = Q_1 = T_2

But Q_1/T_1 represents the increase in entropy ΔS during isothermal expansion

Energy wasted = $T_2 \Delta S$

If T_2 is constant, the amount of energy wasted is proportional to the increase in entropy.

1.20 Third Law of Thermodynamics

With a decrease in temperature, a greater degree of order prevails in any system. If we could cool a system to 0K, the maximum conceivable order would be established in the system and the minimum entropy would correspond to this state. Now, suppose we apply a pressure on the system at 0K. What does happen to the entropy of the system? On the basis of experiments conducted at low temperatures, W. Nernst concluded that at 0K any change in the state of a system takes place without change in the entropy. This is called Nernst's theorem. It is called the third law of thermodynamics. Third law of thermodynamics is sometimes known as the principle of unattainability of absolute zero. It is stated as follows: It is impossible to attain a temperature of 0K.

******Solved examples******

Based on equivalence of heat and work

Ex.1 The height of Niagra falls is 50m. Calculate the difference between the temperature of water at the top and bottom of the fall, if J = 4.2J/cal/kg^{0}C, J = 4.2J/cal

Sol. Given: h = 50m, g = 9.8m/s^2, s = 1kcal

Let m kg of water falls in one second. The potential energy lost in one second is

$$W = mgh = m \times 9.8 \times 50J$$

The lost energy is converted into heat. If Q be the heat produced, then

$$Q = \frac{W}{J} = \frac{m \times 9.8 \times 50}{4.2} = 117\text{mcal} = 0.117\text{m kcal}$$

If this heat causes temperature rise ΔT in water, then

Q = mass × specific heat of water × temperature rise

$$0.117m = m \times 1 \times \Delta T$$

$\Delta T = 0.117^0$ C

Based on work done in different process

Ex.2 If in an isothermal expansion, the volume of 1g mole of gas at 27^{0}C is doubled, calculate the work done in the process (R = 8.3 Jmol^{-1}K^{-1})

Sol. Given: T = 27^{0}C= 300K, R = 8.3 Jmol^{-1}K^{-1}, $V_2 = 2V_1$

The work done in isothermal process is

$$W = 2.3026\ RT\log_{10}\frac{V_2}{V_1}$$

$$= 2.3026 \times 8.3 \times 300\log_{10}\frac{2V}{V}$$

$$= 2.3026 \times 8.3 \times 300 \times 0.3010 = 1725.8\,J$$

Ex.3 A definite mass of a perfect gas is compressed adiabatically to half of its original volume. Determine the resultant pressure if its initial pressure was 1 atmosphere. [Υ = 1.4 and $2^{1.4}$ = 2.64]

Sol. Given: $P_1 = 1$ atmosphere, $V_2 = V_1/2$, $\gamma = 1.4$ and $2^{1.4} = 2.64$
For an adiabatic change

$$P_1V_1^{\gamma} = P_2V_2^{\gamma}$$

Thus , $P_2 = P_1[\frac{V_1}{V_2}]^{\gamma} = 1 \times [\frac{V_1}{V_1/2}]^{1.4} = 1 \times (2)^{1.4} = 2.64$ **atmosphere**

Ex.4 What is the highest possible theoretical efficiency of a heat engine operating with a heat reservoir of furnace gases at 2100⁰C, when cooling water available is at 15⁰C?

Sol. Given: Temperature of furnace, $T_1 = 2100 + 273 = 2373$ K and Temperature of cooling water, $T_2 = 15 + 273$ K $= 288$K

Now, $\eta = 1 - \frac{T_2}{T_1} = 1 - \frac{288}{2373} = 0.878$ or 87.8%

Ex.5 The efficiency of Carnot cycle is 1/6. On reducing the temperature of the sink by 60^0C, the efficiency increases to 1/3. Find the initial and final temperatures which the cycle is working.

Sol. Let T_1 and T_2 be the initial Kelvin temperature of the source and sink respectively

Then the efficiency is given by : $\eta = 1 - \frac{T_2}{T_1} = 1/6$ (i)

When T_2 is decreased to $T_2 - 60^0$C (1^0C = 1K in size). The new efficiency is

$\eta' = 1 - \frac{T_2 - 60}{T_1} = 1/3$ (ii)

From equations (i) and (ii)

$$\frac{T_2}{T_1} = 1 - \frac{1}{6} = \frac{5}{6}$$

And $\frac{T_2-60}{T_1} = 1 - \frac{1}{3} = \frac{2}{3}$

Dividing (ii) by (i) we get

$$\frac{T_2}{T_2-60} = \frac{5}{6}$$

Solving we get $T_2 = 300\text{K} = 27^0\text{C}$

Putting the value of T_2 in (i), we get

$$\frac{300}{T_1} = \frac{1}{6}$$

This gives $T_1 = 360\text{K} = 87^0\text{C}$

The cycle is initially working between 87^0C and 27^0C . finally the temperature of sink is reduced 60^0C , so that the cycle works between 87^0C and -33^0C

Ex.6 A Carnot's refrigerator absorbs heat from water at 0^0C and rejects it at the room temperature 37^0C. Calculate the amount of work required to convert 10kg water at 0^0C into ice at same temperature (latent heat of ice = 3.4×10^5 Jkg⁻¹. Also find the coefficient of performance of the refrigerator.

Sol. Given: m = 10kg, L = 3.4×10^5 Jkg^{-1}, $T_1 = 37^0\text{C} = 310\text{K}$, $T_2 = 0^0\text{C} = 273\text{K}$

We know that for an ideal Carnot's cycle

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2}$$

$T_2 \rightarrow$ lower Temp

$T_1 \rightarrow$ Higher Temp

Or $Q_1 = Q_2\frac{T_1}{T_2}$

Here, $Q_2 = mL = 10 \times 3.4 \times 10^6 = 3.4 \times 10^6\text{J}$

Then $Q_1 = 3.4 \times 10^6 \times \frac{310}{273} = 3.86 \times 10^6\text{J}$

Therefore, work done by the refrigerator

$$W = Q_1 - Q_2 = (3.86 - 3.4) \times 10^6 = 0.46 \times 10^6 = 4.6 \times 10^5 J$$

Coefficient of performance, $\beta = \frac{Q_2}{W} = \frac{3.4\times10^6}{4.6\times10^5} = \frac{34}{4.6} = 7.4$

Imp **Ex.7** A bullet moving with velocity 40m/s falls down after striking a target. If only half of the heat produced is absorbed by the bullet, find rise in its temperature. Specific heat of lead = 120 $Jkg^{-1}degC^{-1}$

Sol. Given: v = 40m/s, s = 120 $Jkg^{-1}degC^{-1}$

Let the mass of the bullet be m kg and its velocity be v.

Then before striking the target, the kinetic energy of the bullet = $\frac{1}{2}mv^2$

On striking the target this is converted into heat

$Q = \frac{1}{2}mv^2$

As the bullet absorbs only half of the heat produced in the impact, heat received by the bullet.

$Q' = 0.5, \quad Q = \frac{1}{2} \times \frac{1}{2}mv^2$

If the increase in the temperature of the bullet is ΔT^0C,

Then $Q' = ms\Delta T$ cal

Hence $ms\Delta T = \frac{1}{2} \times \frac{1}{2}mv^2$

Thus, $\Delta T = \frac{v^2}{4s} = \frac{40^2}{4\times120} = 3.33°C$

Ex.8 The volume of 1g mole of a gas filled in a container at standard pressure $(1 \times 10^5 N/m^2)$and the temperature (0^0C) is $22.4 \times 10^{-3}m^3$. The volume of the gas is reduced to half of its original value by increasing the pressure(i)isothermally (ii) adiabatically. In each case calculate the final pressure of the gas and the amount of work done ($R = 8.3\ Jmol^{-1}K^{-1}, \Upsilon = 1.4$)

Sol. Given , $P_1 = 1 \times 10^5 N/m^2$, $T_1 = 0°C = 273K$, $V_1 = 22.4 \times 10^{-3}m^3$, $V_2 = \frac{1}{2}V_1 = 11.2 \times 10^{-3}m^3$

(i) In an isothermal process

$$P_1V_1 = P_2V_2$$

Or $P_2 = \frac{P_1V_1}{V_2} = 1 \times 10^3 = \frac{V_1}{V_1/2} = 2 \times 10^5 N/m^2$

Work done in the process

$$W = 2.3026\ RT\log_{10}\frac{V_2}{V_1} = 2.3026 \times 8.3 \times 273 \log_{10}\frac{V_1}{2V_1}$$

$$= 2.3026 \times 8.3 \times 273 \log_{10}\frac{1}{2_1} = 2.3026 \times 8.3 \times 273 \times (-0.3010)$$

$$= -1.57 \times 10^3 J$$

(ii) In an adiabatic process

$$P_1V_1{}^{\Upsilon} = P_2V_2{}^{\Upsilon}$$

or $P_2 = P_1[\frac{V_1}{V_2}]^{\Upsilon}$

Final temperature of the gas

$$T_2V_2{}^{\Upsilon-1} = T_1V_1{}^{\Upsilon-1}$$

Or $T_2 = T_1[\frac{V_1}{V_2}]^{\Upsilon-1} = 273 \times [\frac{V_1}{V_1/2}]^{1.4-1} = 273 \times 2^{0.4} = 360.2K\,]$

Work done $W = \frac{R}{\gamma-1}(T_1 - T_2) = \frac{8.73\times(273-360.2)}{1.4-1} = \frac{8.3\times(-87.2)}{0.4}$

$= -1.81 \times 10^3 J$

heat input — work done

Ex.9 A person consumes a diet of 10^4 J/day and spends total energy of 1.2×10^4 J/day. Determine the daily change in the internal energy. If the net energy spent comes from sucrose at the rate of 1.6×10^4 ~~2×10^5 J/kg~~, in how many days will the person reduce his mass by 1 kg?

Sol. Given $dQ = 10^4$ J/day , $dW = 1.2 \times 10^4$ J/day.

From the first law of thermodynamics

$$dU = dQ - dW = 1\times10^4 - 1.2\times10^4 = -2.0\times10^3 J/day$$

This decrease corresponds to the loss of sucrose. Therefore, sucrose lost per day

$$= \frac{2.0\times10^3}{1.6\times10^4} kg = 0.125 \ kg/day$$

Thus, number of days required for the loss of 1kg

$$= \frac{1}{0.125} = 8 \ days$$

Ex.10 At normal temperature (0^0C) and at normal pressure ($1.1.013\times10^6 \ N/m^2$) when one gm of water freezes its volume increases by 0.09 /cm³. Calculate the change in its internal energy (latent heat of melting ice = 80 cals/gm, 1cal = 4.2 J)

Sol. 1.$P = 1.013\times10^6 \frac{N}{m^2} = 1.013\times10^6 \frac{dynes}{cm^2}$, m = 1gm, V_1=1cm³, V_2=0.091 cm³

Work done $dW = PdV = -1.013\times10^6\times0.091 = -9.2\times10^4 erg$

$$dQ = 80cal = 80\times4.2\times10^7 erg = 336\times10^7 erg$$

According to the first law of thermodynamics

$$dQ = dU + dW$$

$$dU = dQ - dW$$

$$dU = 336\times10^7 + 9.2\times10^4 = (336 + 0.0092)\times10^7 erg = 336 J$$

Ex.11 1 mole of a gas at 127^0C expands isothermally until its volume is doubled. Calculate the work done.

Sol. In isothermal process,

$\boxed{1 \ erg = 10^{-7} \ Joule}$

$W = RT\ln\frac{V_2}{V_1}$ $[V_2 = 2V_1]$

$R = 8.3\times10^7$, T = 127+273=400K $\boxed{2310.1 J}$

So $W = 22.9\times10^9 ergs$

Ex.12 A Carnot's engine whose low temperature reservoirs at 7^0C has an efficiency of 40%. To increase the efficiency to 50% by how many degree should the temperature of the source be increased?

Sol. The efficiency is given by : $\eta = 1 - \frac{T_2}{T_1}$

Or $\eta' = 1 - \frac{T_2}{T_1'}$, $T_2 = 280K$

T_1= 466.67K, $T_1' = 560K$

ΔT = 560 - 466.67 = 93.33K = -179.67^0C

Ex.13 A Carnot's engine working between temperatures 100^0C and 0^0C absorbs 10^4 calories of heat at higher temperature. What will be the work done when the cycle of operation is over.

Sol. $\eta = 1 - \frac{T_2}{T_1} = \frac{W}{\text{heat absorbed}}$

$W = 1.126\times10^4 J$

Ex.14 Calculate the change in entropy when 2g of ice melts into water at the temperature 0^0C. Latent heat of ice = 80cal/g.

Sol. $dQ = T\,dS$

$dS = dQ/T = 160/273 = 0.58cal/K = 0.58\ cal/^0C$

*****Review Questions*****

Based on laws of thermodynamics

1.What are fundamental ideals of thermodynamics?

2.Discuss continuum model of thermodynamics

3.Define the following: (a) system (b) state (c) equilibrium (d) process.

4.Explain how first law of thermodynamics leads to concept of internal energy ?

5.Why does the temperature of a gas drops when it is subjected to adiabatic expansion? Explain

6.State First law of thermodynamics and prove that internal energy is a thermodynamical variable.

7.Write notes on following (a) Zeroth law of thermodynamics and concept of temperature (b) Zeroth law of thermodynamics (c) Concept of internal energy (d) First law of thermodynamics

8.Explain in short first law of thermodynamics.

9.What do you mean by indicator diagram? Draw indicator diagram for isochoric and isobaric processes.

10.What do you understand by internal energy of a system?

11.What is zeroth law of thermodynamics? State its importance?

12.State first law of thermodynamics. What are its limitations ?

13.State the zeroth law of thermodynamics. How is mercury in thermometer able to find temperature of a body using the zeroth law of thermodynamics ?

14.Explain the first law of thermodynamics for the closed system undergoing a cyclic change.

15.Write the zeroth law of thermodynamics. On this basis explain the concept of temperature. Explain the zeroth law of thermodynamics.

16.What is the concept of temperature? Define temperature.

17.Prove that heat and work both are the path function.

18.State first law of thermodynamics.

19.Show that for a cyclic process the heat supplied to a system is equal to the work done by the system. Define reversible and irreversible process with one example of each process.

20.State zeroth law of thermodynamics and explain the concept of temperature on the basis of this law.

21.State the first law of thermodynamics. Discuss its significance.

22.Calculate work done in an adiabatic expansion of a perfect gas.

23. State first law of thermodynamics. and show that heat and work are path functions but their difference is a point function.

24.Write down the zeroth law of thermodynamics. Explain how it is introduces the temperature of a system as its function of state.

25.State the first law of thermodynamics and use it to derive a relation between the volume and temperature of a perfect gas undergoing an adiabatic change.

26.Explain the first law of thermodynamics. Explain the latent heat on the basis of it.

27.State two statements of second law of thermodynamics and show their equivalence.

28. State second law of thermodynamics.

29.Show that work is a path function and not a property.

30.State the first law for a closed system undergoing a change of state.

31.What is thermodynamics? State the first, second and third laws of thermodynamics and discuss their significance

32.Why the second law is called a directional law of nature?

33.Give the Nernst statement of the third law of thermodynamics?

34.Give the Kelvin-Planck statement of the second law.

35.Give the Clausius statement of the second law.

Based on Carnot Cycle, Carnot Engine, Entropy

36.Derive the expression for the efficiency of a Carnot engine directly from a T-S diagram.

37.What are the reversible processes?

38.Write the Kelvin-Planck and Clausius statements of the second law of thermodynamics.

39.Between two given temperatures no ordinary engine can be more efficient than the Carnot engine and all Carnot's engines are equally efficient. Prove the statement.

40.State the essential conditions for a process to be reversible.

41.What do you understand by entropy? State the second law of thermodynamics in the entropy.

42.What is Carnot's theorem? Prove it.

43.Explain need of second law of thermodynamics. State it's both statements and show the equivalence.

44.Prove that the efficiency of a Carnot's engine depends only upon the two temperature between which it works.

45.Differentiate between reversible and irreversible processes.

46.What is a reversible process? A reversible process should not leave any evidence to show that the process had ever occurred. Explain.

47.All spontaneous processes are irreversible. Explain.

48.Distinguish between a reversible and an irreversible process. Illustrate your answer by some examples.

49.A Carnot's engine and refrigerator work between same temperatures T_1K and T_2K. Write expressions for the efficiency η of Carnot's engine and coefficient of performance β, the refrigerator and inter relationship between η and β.

50.Write the expression for efficiency of a Carnot's reversible engine.

51.Establish a relation between the efficiency η of an ideal Carnot engine and the coefficient of performance β of an ideal refrigerator working between temperatures T_1K and T_2K

52.Explain second law of thermodynamics.

53.What is significance of second law of thermodynamics?

54.What is physical significance of entropy?

55.The entropy of a substance is a unique function of its state. Explain.

56.Prove that the integral $\int_1^2 \frac{dQ}{T}$ does not depend on path for a reversible process.

57.Define and explain entropy. Explain why unavailable energy in the universe tends to increase?

58. Show that entropy remains constant in a reversible process whether it increases for an irreversible process.

59.Define entropy and explain its physical significance?

60.What is the absolute scale to temperature and how has it been derived? Explain clearly why the scale is called absolute and why the zero of this scale is considered to the lowest temperature possible?

61.State and Prove Carnot's theorem.

62.What is a cyclic heat engine?

63. Define the thermal efficiency of a heat engine cycle. Can this be 100%?

64. What is a Carnot Cycle? What are the four processes which constitute the cycle?

65.Describe the different operations involved in a Carnot's cycle. Derive the efficiency of a Carnot engine in terms of source and sink temperatures.

66.What is a reversed heat engine?

67.Show that the efficiency of a reversible engine operating between two given constant temperatures is the maximum.

68.How does the efficiency of a reversible engine vary as the source and sink temperatures is varied? When does the efficiency become 100%?

69.What do you understand .by the entropy principle?

70.Show that the transfer of heat through a finite temperature difference is irreversible?

71.Show that the adiabatic mixing of two fluids is irreversible.

72.What is the maximum work obtainable from two finite bodies at temperatures T_1 and T_2?

****

Unit - II

Waves and Oscillations: Wave motion, simple harmonic motion, wave equation, superposition principle. Introduction to Electromagnetic Theory, Maxwell's equations, Work done by the electromagnetic field, Poynting's theorem, Momentum, Angular momentum in electromagnetic fields, Electromagnetic waves: the wave equation, plane electromagnetic waves, energy carried by electromagnetic waves

Chapter 2

Wave and Oscillations

Introduction

Before wave motion, we have to understand oscillatory motion which is a kind of periodic motion. Therefore, we need to understand first the periodic motion.

Periodic Motion: A motion that repeats itself after a certain period of time, is called periodic motion. Repetition of the motion may be in a circle or a back-and-forth motion about a fixed mean point.

Oscillatory Motion: When the motion repeats itself in a back-and-forth manner about a fixed point, it is termed as oscillatory motion. It may also be called vibratory motion if the frequency of repetitions is quite high.

From the above discussion, it is clear that oscillatory motion should have the following characteristics.

Frequency: The numbers of repetitions per time, done by the oscillatory body, is called the frequency (f), of the periodic motion.

Time Period: Time taken to complete, one oscillation, is called the time period (T).

Here, we can have $T = \frac{1}{f}$

Amplitude: The maximum distance travelled on either side of the fixed point, is called amplitude.

2.1 Simple Harmonic Oscillation

It is a kind of oscillatory motion in which the motion is in a straight path on both the sides of the fixed point. If we look at the system of S.H.M., as a whole, it can be said that a force, called restoring force is needed to perform the S.H.M and it always follows Hook's Law, i.e.

$$F = -kx,$$

Here, x is the displacement from the mean position and k is the force constant.

2.1.1 Equation of S. H. M. and its Solution

Consider a particle of mass 'm' executing S.H.M along x axis about its mean position such that the displacement from the mean position be x. Then, according to the condition to perform S.H.M, restoring force F will act on the particle to oppose its motion and is proportional to the displacement. Therefore;

$$F \propto -x \text{ or } F = -kx \tag{1}$$

Here, k is the constant of proportionality and is known as force constant or force per unit displacement.
Now, according to the Newton's second law of motion, **"the acceleration produced, is defined as the force applied per unit mass of a moving object"** i.e.

$$\frac{d^2x}{dt^2} = \frac{F}{m}$$

Or we can write, Force = mass x acceleration i.e.

$$F = m\frac{d^2x}{dt^2} \qquad (2)$$

From (1) and (2)

$$m\frac{d^2x}{dt^2} = -kx \quad \text{or} \qquad \frac{d^2x}{dt^2} = -\frac{k}{m}x$$

Putting $\boxed{\frac{k}{m} = \omega^2}$ (3)

We have

$$\frac{d^2x}{dt^2} + \omega^2 x = 0 \qquad (4)$$

This is a second order differential equation and its solution can be written as;

$$x = A\sin(\omega t + \delta) \qquad (5)$$

The equation (5), gives the instantaneous displacement of the particle, executing S.H.M at any instant of time. A is the maximum displacement of the particle and is called the displacement amplitude and $(\omega t + \delta)$ is called instantaneous phase angle.

2.2 Ideal Simple Harmonic Motion

The ideal S.H.M. can be defined with the help of a particle moving in a circle. "If a particle is moving on a circular path and a perpendicular is drawn on any diameter of the circle, then the motion of foot of the perpendicular about the centre of the circle is called ideal S.H.M" as shown in Fig. 2.1 The equation for instantaneous displacement can be obtained very easily using Fig. 2.1

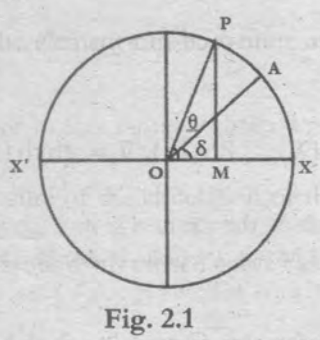


Fig. 2.1

Let a particle, start its motion on a circular track with the initial phase XOA as δ. Let us consider any arbitrary position P of the particle such that angle AOP is θ. Let PM is a perpendicular drawn on the diameter XX′. As the particle moves on the circular path and completes a circle, the point M i.e., the foot of the perpendicular, moves on the straight path XX′, about the mean point O and also complete one oscillation. The motion of this point M, is an oscillatory motion about the centre of the circle and it is in perfect straight path, termed as ideal simple harmonic motion. So "the motion of the foot of the perpendicular, drawn on any diameter from a particle, moving on a circular track is called ideal simple harmonic motion". The amplitude of the oscillation A is equal to the radius of the circle OX. Now from the diagram, we can write, θ

$$x = A\sin(\theta + \delta)$$

Where $\theta = \omega t$, and ω is the angular frequency, i.e., angular distance per unit time.

Or $\quad \omega = \frac{\theta}{t} = \frac{2\pi}{T} = 2\pi f$

From equation (3), we have the angular frequency $\omega = \sqrt{\frac{k}{m}}$, so the time period and the frequency of S.H.M will be written as;

$$T = 2\pi\sqrt{\frac{m}{k}}$$

And,
$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

2.2.1 Velocity and Acceleration in S.H.M

We can find the velocity of the particle executing S.H.M. from the expression of the displacement.

$$x = A\sin(\omega t + \delta) \tag{6}$$

Differentiating it w.r.t. time we get;

$$v = \frac{dx}{dt} = A\omega\cos(\omega t + \delta)$$

$$= A\omega\sqrt{1 - \sin^2}(\omega t + \delta) = \omega\sqrt{A^2 - A^2\sin^2(\omega t + \delta)}$$

$$\boxed{v = \omega\sqrt{A^2 - x^2}} \tag{7}$$

This is the expression for the velocity of the particle at any point which has a displacement x from its mean position. It is well known that particle has maximum velocity at the mean position and minimum at the extreme point, due to obvious reasons. So, the maximum velocity can be obtained by putting x = 0 and minimum at x = A. i.e.

$v_{max} = A\omega$ (at mean position)

$v_{min} = 0$ (at extreme position)

Again, differentiating equation (7) w.r.t time, we get acceleration as;

$$\alpha = \frac{dv}{dt} = -A\omega^2\sin(\omega t + \delta)$$

$$\alpha = -\omega^2 x \tag{8}$$

The above equation gives acceleration of the oscillating particle at any displacement. From equation (8), it can be seen that acceleration is proportional to the displacement and in opposite direction of the motion. This is a standard characteristic of a simple harmonic motion. It is clear from equation (8) that for the maximum acceleration x is equal to the amplitude A.

So, the maximum acceleration i.e., $\alpha_{max} = \omega^2 A$ (at the extreme position)

And the minimum acceleration i.e., $\alpha_{min} = 0$ (at the mean position)

2.3 Energy of a Simple Harmonic Oscillator

A simple harmonic oscillator possesses both the kinetic energy and the potential energy. Here it can also be seen that in ideal condition, the law of conservation of energy also holds good during the complete oscillation of the particle.

The potential energy is given by the amount of work done stored in displacing the particle from the position 0 to x by applying some force. Thus,

$$P.E. = \int \vec{F}.\vec{dx} = \int_0^x kx\,dx = \boxed{\frac{1}{2}kx^2} \tag{9}$$

Putting the values of k and x, the instantaneous potential energy at any position is;

$$\boxed{\mathrm{P.E.} = \frac{1}{2} m\omega^2 A^2 \sin^2(\omega t + \delta)} \qquad (10)$$

And the corresponding kinetic energy of the harmonic oscillator, at any instance, is given as

$$\mathrm{K.E.} = \frac{1}{2} m v^2 = \frac{1}{2} m \left[\frac{dx}{dt}\right]^2$$
$$= \frac{1}{2} m\omega^2 A^2 \cos^2(\omega t + \delta)$$

So, the total energy E of the oscillator at any displacement from the mean position can be obtained as

$$E = \mathrm{K.E} + \mathrm{P.E}$$
$$= \frac{1}{2} m\omega^2 A^2 \cos^2(\omega t + \delta) + \frac{1}{2} m\omega^2 A^2 \sin^2(\omega t + \delta)$$
$$E = \frac{1}{2} m\omega^2 A^2 \qquad (11)$$

Hence, the total energy comes out to be constant and independent of the position of the particle. It is obvious that the maximum possible value of K.E and P.E would be same as $\frac{1}{2} m\omega^2 A^2$. The K.E and P.E of the harmonic oscillator can also be calculated with the passage of time, in place of displacement. If the oscillation starts at t = 0 at its extreme position, then;

The average P.E of the simple harmonic oscillator over a cycle, is given as:

$$< \mathrm{P.E} > = \frac{1}{T}\int_0^T \frac{1}{2} k x^2 dt$$
$$= \frac{\int_0^T \frac{1}{2} k A^2 \sin^2(\omega t + \delta) dt}{T}$$
$$= \frac{1}{2}\frac{\int_0^T m\omega^2 A^2 \sin^2(\omega t + \delta) dt}{T}$$
$$= \frac{1}{2} m\omega^2 A^2 \frac{\int_0^T \sin^2(\omega t + \delta) dt}{T}$$
$$< \mathrm{P.E} > = \frac{1}{4} m\omega^2 A^2 \qquad \text{as } \left[\frac{\int_0^T \sin^2(\omega t + \delta) dt}{T} = 1/2\right] \quad (12)$$

And similarly, the average kinetic energy for one complete cycle is;

$$< \mathrm{K.E} > = \frac{1}{T}\int_0^T \frac{1}{2} m \left(\frac{dx}{dt}\right)^2 dt$$
$$= \frac{\int_0^T \frac{1}{2} m A^2 \omega^2 \cos^2(\omega t + \delta)\, dt}{T}$$
$$= \frac{1}{2} m\omega^2 A^2 \frac{\int_0^T \cos^2(\omega t + \delta)\, dt}{T}$$
$$< \mathrm{K.E} > = \frac{1}{4} m\omega^2 A^2$$
$$\Rightarrow < \mathrm{K.E} > = < \mathrm{P.E} > = \frac{1}{4} m\omega^2 A^2$$
$$= \frac{1}{4} m\omega^2 A^2 \qquad (13)$$

From equations (12) and (13), it can be seen that the total energy is equally divided into kinetic as well as potential energy. From here, it is also clear that the total energy of a simple harmonic oscillator is always conserved.

2.4 Examples of S.H.M

2.4.1 Simple Pendulum

A point mass attached to a massless, inextensible string is called a simple pendulum as in Fig.2.2. Let us find out various physical parameters for a simple pendulum.

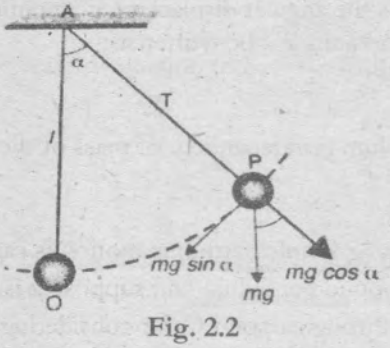


Fig. 2.2

Let a particle of mass m, at point P, is under the action of various forces:

(i) The weight mg of the bob acting vertically downward

(ii) Tension 'T', in the string along PA.

From the diagram, it can be seen that the tension is opposed by the radial component of weight, $mg\cos\alpha$, therefore, the force $T - mg\cos\alpha$ provides centripetal force for circular arc and the tangential component $mg\sin\alpha$ tends to bring the bob back to its initial position 'o'. Here, $mg\sin\alpha$, is known as restoring force and is responsible for performing oscillations of the particle. So,

$$F = -mg\sin\alpha$$

The negative sign indicates that the acceleration and the displacement are oppositely directed. If $\frac{d^2x}{dt^2}$ be the acceleration at any time t, in the direction of increasing x, then the force must be;

$$F = m\frac{d^2x}{dt^2}, \text{ so}$$

$$m\frac{d^2x}{dt^2} = -mg\sin\alpha$$

$$\frac{d^2x}{dt^2} = -g\sin\alpha$$

For the condition to perform S.H.M, the angle α should be small, so the distance x (arc) can be written as;

$x = l\alpha,$ (where, l is the radius of the circular path, i.e., length of the thread)

On differentiating two times, it gives;

$$\frac{d^2x}{dt^2} = l\frac{d^2\alpha}{dt^2}$$

$$l\frac{d^2\alpha}{dt^2} = -g\alpha \quad (\sin\alpha \approx \alpha)$$

$$\frac{d^2\alpha}{dt^2} = -\frac{g}{l}\alpha$$

$$\frac{d^2\alpha}{dt^2} + \frac{g}{l}\alpha = 0$$

Considering, $\omega^2 = \frac{g}{l}$, we have;

$$\frac{d^2\alpha}{dt^2} + \omega^2\alpha = 0 \tag{14}$$

The equation (14), is a second order differential equation and is the equation of the motion of simple pendulum, with solution as;

$$\boxed{\alpha = \alpha_0 \sin(\omega t + \delta)}$$

Here, δ is the initial phase and α_0 is the angular displacement amplitude.

The time period for the simple pendulum will be written as;

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{l}{g}} \tag{15}$$

The time period of a simple pendulum is independent of mass of the bob.

2.4.2 Compound Pendulum

An object of laminar shape, executing simple harmonic motion is called a compound pendulum. In Fig. 2.3, let 'm' is the mass of the compound pendulum and suppose it is making small oscillations under the influence of gravity, about an axis through a point O for considering the motion as S.H.M. Let G be the centre of gravity of the compound pendulum and l be the distance OG.

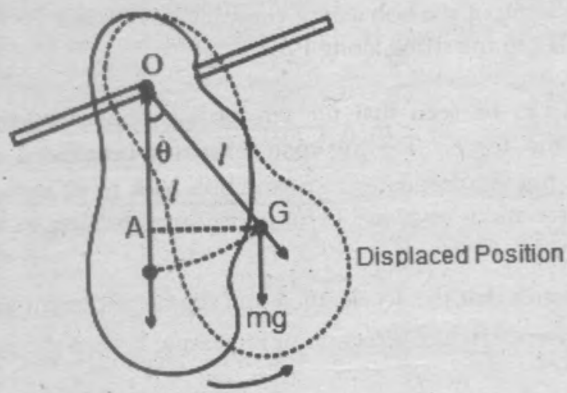


Fig. 2.3

The pendulum is displaced from its mean position so that OG makes a small angle θ with vertical line OA. In the displaced position, the weight mg of the pendulum, acting vertically downwards produces a torque which tends to bring the pendulum to its initial position and is given by;

$$\tau = mg \times GA = mgl\sin\theta$$ (force × length of arm)

As we know the torque can also be written as

$$\tau = -I\alpha$$

Where I is the moment of inertia of the pendulum about O and $\alpha = \frac{d^2\theta}{dt^2}$ is the angular acceleration. So, we have;

$$\boxed{mgl\sin\theta = -I\left(\frac{d^2\theta}{dt^2}\right)}$$

$$I\frac{d^2\theta}{dt^2} + mgl\sin\theta = 0$$

Here θ is small for obvious reason, So, $\sin\theta = \theta$. Thus,

$$I\frac{d^2\theta}{dt^2} + mgl\,\theta = 0$$

$$\frac{d^2\theta}{dt^2} + \frac{mgl}{I}\theta = 0 \qquad (16)$$

Considering, $\omega^2 = \frac{mgl}{I}$, equation (16) can be written as

$\frac{d^2\alpha}{dt^2} + \omega^2\alpha = 0$, which is again a standard equation of S.H.M., with the time period

$$T = \frac{2\pi}{\omega}$$

$$T = 2\pi\sqrt{\frac{I}{mgl}}$$

If K is the radius of gyration of the pendulum about a parallel axis through G then using, parallel axis theorem, the moment of inertia is written as;

$$I = mK^2 + ml^2$$

Thus;

$$T = 2\pi\sqrt{\frac{K^2 + l^2}{gl}}$$

$$T = 2\pi\sqrt{\frac{\frac{K^2}{l}+l}{g}} \qquad (17)$$

Now if we compare equations (15) and (17), we see that both the equations are similar, if we put $\frac{K^2}{l} + l$ = L, i.e. the formula for the time period for a compound pendulum, takes the form as;

$$T = 2\pi\sqrt{\frac{L}{g}} \qquad (18)$$

Here, $L = \frac{K^2}{l} + l$ is the equivalent length of the compound pendulum. It means, if we are able to find the equivalent length 'L', of an oscillating compound pendulum, then the compound pendulum can be considered as a simple pendulum with length of the thread as 'L'. The equation (18), is a very important equation to find the earth's gravitational acceleration, g, at any place without using sophisticated instrumentation.

2.5 Applications of Simple Harmonic Oscillations

As discussed above, besides finding acceleration due to gravity, there are enormous practical examples that can be considered as the direct applications of simple harmonic motion, but due to limitations to our topics of concern. Here we will discuss a few ideal cases, where no damped or forced oscillations are involved.

2.5.1 Oscillations of L-C circuit

If a charged capacitor is attached to an inductor, the electrical energy stored in the capacitor, is converted into the magnetic energy of the inductor and again the magnetic energy converted back into the electrical energy of the capacitor due to the oscillations of the charge of the capacitor in the circuit. This conversion goes on for infinite time if there is no dissipation of energy during this process. Practically it is not possible, so this may be treated as an application of an ideal case of simple harmonic oscillation.

Suppose a charged capacitor C, is connected with an inductor L, Fig. 2.4. Let q_0 is the total charge stored in the capacitor of capacitance 'C', then the potential difference between the plates of the capacitor will be $\frac{q_0}{C}$.

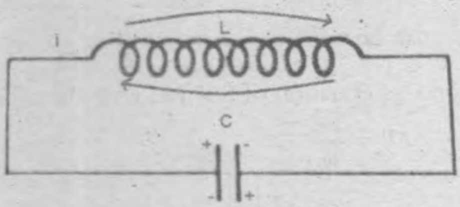


Fig. 2.4.

Now if the circuit is closed, the charge from positive plate of the capacitor starts flowing towards the negative plate of the capacitor, through the inductor 'L'. Let 'q' be the instantaneous charge and 'i' is the instantaneous current in the circuit, then the e.m.f. induced between the ends of the inductor will be $L\frac{di}{dt}$. As there is no external source, then in a close circuit, the total potential will be zero i.e.

$\frac{q}{C} + L\frac{di}{dt} = 0$, or

$\frac{q}{C} + L\frac{d^2q}{dt^2} = 0$, or

$\frac{d^2q}{dt^2} + \frac{1}{LC}q = 0$, or

$$\frac{d^2q}{dt^2} + \omega^2 q = 0 \qquad (19)$$

The equation (19) is a standard differential equation of simple harmonic motion. Here, $\frac{1}{\sqrt{LC}}$ is the angular frequency of the oscillations and charge 'q', is the physical variable of the motion, having amplitude as q_0. The frequency is given as,

$\nu = \frac{\omega}{2\pi}$. i.e.

$$\nu = \frac{1}{2\pi\sqrt{LC}}$$

And the instantaneous charge is calculated by the solution of the equation (19), as;

$q = q_0 \sin(\omega t + \phi)$

And the instantaneous current will be;

$i = \frac{dq}{dt} = \omega q_0 \cos(\omega t + \phi)$

The total energy in the circuit either electrical or magnetic will be

$\frac{1}{2}CV^2 = \frac{1}{2}\frac{q_0^2}{C}$, (Purely electrical)

$\frac{1}{2}CV^2 = \frac{1}{2}Li_0^2$ (Purely magnetic)

Or

$\frac{1}{2}\frac{q^2}{C} + \frac{1}{2}Li^2$ (Instantaneous energy, electrical and magnetic)

2.6 Wave Motion

When a particle in a material medium executes S.H.M. then due to the inter molecular force the neighbouring particles of the medium, also start oscillating simple harmonically, either in the direction of motion of the particle (longitudinally) or in the perpendicular direction of the motion of the source particle (transversely) or in both the directions. It means the energy of the oscillating particle gets transferred to its neighbouring particle without the actual movement of the particle of the medium. This process of movement of the energy is termed as wave motion and this transfer/movement of energy/disturbance, in the medium is called a wave. It may be progressing through the medium, called progressive wave or redistributed among the particle in the still medium, called as standing wave. The origin of energy transfer in the medium is the inter-molecular force between the particles of the medium (in case of a mechanical wave). The mechanism of propagation of an electromagnetic wave is different from the mechanical wave.

2.7 Wave Equation

As the wave motion through the medium is associated with the oscillation of the particle at the source which is oscillating simple harmonically or otherwise, the wave equation can be derived with the help of the equation of simple harmonic motion of the source particle. As the motion of the particle at the source is governed by an oscillating external force, in the same manner the wave motion is governed by the oscillation of the particle at the source. It can be seen that the particle at the source oscillates with the frequency of governing force. Similarly, the frequency of the wave is also equal to the frequency of the particle at the source and again the wave motion is simple harmonic if the particle at the source is oscillating simple harmonically.

Now, let us write the equation of simple harmonic oscillations, performed by the particle as;

$$y = A\sin(\omega t + \delta)$$

Here y is the instantaneous displacement of the particle from the mean position, oscillating simple harmonically along y-axis. A is the displacement amplitude. ω and δ are the frequency and initial phase of the particle, respectively. Now starting the motion with initial phase as zero, the displacement of the particle at $x = 0$, will be;

$$y = A\sin\omega t$$

So, the displacement of the particle at position x from the mean position and at time 't', will be;

$$y = A\sin\omega\left(t - \frac{x}{v}\right) = A\sin\left(\omega t - \frac{\omega x}{v}\right) \tag{20}$$

The equation (20) may be called as an equation of a wave which is associated with the displacement of the particle at the source. Here v is the velocity of the wave and x is the displacement of the wave from the origin and $\frac{\omega x}{v}$ is the phase w.r.t. t., the origin. From here it can also be seen that after a certain distance say X, the phase will repeat itself after one or more than one time periods. Then the equation (20) can be written as;

$$y = A\sin\omega\left(t - \frac{x}{v}\right) = A\sin\omega\left(t - \frac{x+X}{v}\right)$$

$$\frac{dy}{dt} = A\omega\cos\omega\left(t - \frac{x}{v}\right) = A\omega\cos\omega\left(t - \frac{x+X}{v}\right)$$

From here we can write

$$\omega\left(t - \frac{x}{v}\right) = \omega\left(t - \frac{x+X}{v}\right) + 2n\pi$$

Or

$$\boxed{X = \frac{v}{\omega}2n\pi}$$

For two consecutive point of same phase, $n = 1$, so

$$X = \frac{2\pi v}{\omega} = vT \tag{21}$$

Equation (21) gives the distance travelled by the wave in a complete time period T. This length is called wavelength and denoted by λ. The equation can be re-written in its well-known form, i.e.

$$v = \nu\lambda \tag{22}$$

The other forms of equations of wave are

$y = A\sin(\omega t - \omega\frac{x}{v}) \Rightarrow y = A\sin(\omega t - kx)$,

Where, $k = \frac{w}{v} = \frac{2\pi\nu}{v} = \frac{2\pi}{\lambda}$ is called wave number.

Also, another form of equation is;

$y = A\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)$,

2.8 Particle Velocity and Wave Velocity

The particle velocity of the particle, executing S.H.M at the source, is given by

$$\frac{dy}{dt} = A\omega\cos\omega\left(t - \frac{x}{v}\right) \quad (23)$$

Now, differentiating (23) w.r.t. x, we have

$$\frac{dy}{dx} = -\frac{A\omega}{v}\cos\omega\left(t - \frac{x}{v}\right) \quad (24)$$

From equations (23) and (24), the particle velocity u, can be written as

$$u = \frac{dy}{dt} = v\frac{dy}{dx} \quad \text{(Ignoring -ve sign)} \quad (25)$$

Equation (25) relates the particle velocity to the velocity of the wave associated with the particle, executing S.H.M. at the source.

2.9 General Form of Equation of Wave

Differentiating equations (23) w.r.t. t and (24) w.r.t. x, respectively, we get

$$\frac{d^2y}{dt^2} = A\omega^2\sin\omega\left(t - \frac{x}{v}\right) \quad (26)$$

$$\frac{d^2y}{dx^2} = \frac{A\omega^2}{v^2}\sin\omega\left(t - \frac{x}{v}\right) \quad (27)$$

From equation (26) and (27) we can write;

$$\boxed{\frac{d^2y}{dx^2} = \frac{1}{v^2}\frac{d^2y}{dt^2}} \quad (28)$$

This is called general wave equation in differential form.

2.10 Superposition Principle for Waves

2.10.1 Waves in Same Direction (Progressive Wave)

"When two or more than two waves superimpose at a point in the medium, the instantaneous amplitude of the resultant waves is given by the vector sum of all the instantaneous amplitudes of all the waves". This is called superposition principle of waves. Hence, if there are 'n' number of waves having amplitudes $y_1, y_2, y_3, \ldots\ldots\ldots\ldots y_n$, then;

$y = y_1 + y_2 + y_3 + \ldots\ldots\ldots\ldots y_n$

Here, y is the instantaneous amplitude of resultant wave due to the superposition of the n waves.

For simplicity let us take the case of two waves only. Let

$y_1 = A_1\sin(\omega t + kx)$ and

$y_2 = A_2\sin(\omega t + kx + \emptyset)$,

According to the principle of superposition, the resultant will be;

$$y = A_1\sin(\omega t + kx) + A_2\sin(\omega t + kx + \emptyset) \quad (29)$$

On solving equation (29) we have;

$$y = A\sin(\omega t + kx + \delta) \quad (30)$$

Where A is the amplitude of the resultant wave and δ is the phase of the resultant wave, given as;

$y = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\cos\phi}$

and

$$\tan \delta = \frac{A_2 \sin\phi}{A_1 + A_2 \cos \phi}$$

From equation (30), it is clear that the resultant of the superposition of two simple harmonic waves travelling in the same direction will again be simple harmonic of same frequency as that of the source waves with some phase angle with the source.

2.10.2 Waves in Opposite Direction (Standing Waves)

Let us take two simple harmonic waves travelling in opposite direction to each other. i.e.

$$y_1 = A_1 \sin(\omega t + kx)$$

$$y_2 = A_2 \sin(\omega t - kx + \emptyset)$$

Using superposition principle, the resultant amplitude will be given as;

$$y = A \sin(\omega t + kx) + A \sin(\omega t - kx)$$

Here, phase difference $\emptyset$ is taken as zero, and the amplitude is taken same as a special case.

On simplification we have,

$$y = (2A\cos kx) \sin \omega t \qquad (31)$$

The equation (31) is the equation of standing wave, in which the amplitude i.e., $2A\cos kx$, of resultant waves, also varies sinusoidally. The points of maximum amplitude are called antinode and the points of minimum amplitude are termed as nodes.

sin graph.

2.11 Category of waves

Fundamentally, waves can be divided into two categories. One, which needs a material medium, called mechanical waves and the other which do not need any material medium for their propagation, are called electromagnetic waves.

2.12 Electromagnetic Field Theory

Introduction

A static charge produces static electric field around it and the theory is called electrostatic field theory. The theory behind uniformly moving charge is called magnetostatic field theory. But if the charge is accelerating or oscillating, the theory, which explain this phenomenon, is termed as electromagnetic field theory. An accelerating or oscillating charge produces oscillating electric field and the oscillating electric field produces, oscillating magnetic field. We have the equation of oscillating electric field as;

$$E_y = E_{y_0} \sin(\omega t + kx) \qquad (32)$$

And the equation of corresponding magnetic field will be

$$B_z = B_{z_0} \sin(\omega t + kx) \qquad (33)$$

Later in this chapter, it will be proved that the in case of an electromagnetic wave, the oscillations of electric and magnetic field vectors are mutually perpendicular to each other and also perpendicular to the propagation of the wave.

Maxwell derived and developed the fundamental equations of electromagnetic field theory and ultimately the theory of electromagnetic waves. He also predicted that light is an electromagnetic wave having electric and magnetic field vectors, oscillating perpendicular to each other and also perpendicular to the direction of propagation of electromagnetic energy.

2.13 Maxwell's Equations

A set of four equations of electromagnetism (with some modifications), are called the Maxwell's equations of electromagnetic field theory. Maxwell rigorously analysed four fundamental equations of electromagnetism, i.e., of electric field, magnetic field and electromagnetic induction and came to the conclusion that these equations are interlinked although, they are written separately for electrostatics, magnetostatics or electromagnetic induction. After doing a lot of mathematical operations on these equations, he also derived nearly all mathematical formulations for electromagnetic field theory. Later, this theory led to the theory of electromagnetic waves. Heinrich Hertz was the first scientist who had successfully generated the electromagnetic waves.

After doing mathematical operations on the four fundamental equations of electrostatics, magnetism and electromagnetism, Maxwell found very interesting physical interpretation of these equations in integral as well as in the differential form. He concluded that these equations are not ideal rather they are interlinked to each other and lead to the fundamental foundation of the electromagnetic wave theory.

These four equations of electromagnetism can be converted into two, using the concept of electromagnetic potential. In relativistic physics and using the concept of electromagnetic four-vector potential, there comes out to be only one equation of electromagnetism. Later, this will be explained in detail, in the chapter of "special theory of relativity", under the topic of 'invariance of Maxwell's equations under Lorentz's Transformation'.

2.13.1 Modifications of the Equations of Electromagnetism

Here are the four equations of electromagnetism as;

(i) $\oint E.\,ds = \frac{q}{\epsilon_0}$ — Gauss's Law of Electrostatics

(ii) $\oint B.\,ds = 0$ — Gauss's Law of Magnetostatics

(iii) $\xi = -\frac{\partial \phi_B}{\partial t}$ — Faraday's Law of electromagnetic Induction

(iv) $\oint B.\,dl = \mu_0\, I$ — Ampere's Circuital Law

Initially, Maxwell checked the consistency of each equation in every situation and eventually found that the Ampere's circuital law to be inconsistent with an AC circuit, if a capacitor is connected into an AC circuit. Then, he modified the equation, made it consistent and put all the four equations of electromagnetism, under one umbrella. All together they are called Maxwell's equations of electromagnetism.

2.14 Inconsistency of Ampere's Circuital Law — Maxwell's Displacement Current

According to the Ampere's circuital law, the line integral of the magnetic induction B around any closed path around a current carrying conductor is μ_0 times the current flowing through the conductor and written as;

$$\oint B.\,dl = \mu_0\, I \tag{34}$$

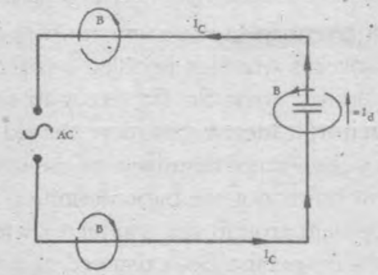


Fig.2.5

As there is no current flowing between the plates of the capacitor, so in the loop over the free space between the plates of the capacitor, the Ampere's law gives;

$$\oint B.\,dl = \mu_0 I_C = 0 \quad (35)$$

So, from (35), it is clear that Ampere's circuital law is not consistent for a circuit in which a capacitor is maintained with alternating electric field, as there is not current flowing in the free space between the plates of the capacitor, Fig. 2.5. To remove this inconsistency, the concept of Maxwell's displacement current was introduced. The expression for displacement current can be derived as follows;

Let an electric circuit contains a capacitor, connected to an AC source. For any instantaneous value of surface charge density σ on the plates, the electric field between the plates is given by;

$E = \frac{\sigma}{\epsilon_0}$, where $\sigma = \frac{q}{A}$,

q is the instantaneous charge and A is the area of the plates and ϵ_0 is the permittivity of the free space between the plates.

Thus, $E = \frac{q}{\epsilon_0 A}$

As the capacitor is connected to an AC source, the time variation of the electric field between the plates is given by;

$$\frac{\partial E}{\partial t} = \frac{1}{\epsilon_0 A}\frac{\partial q}{\partial t},$$

Or

$$\frac{\partial q}{\partial t} = \epsilon_0 A \, \frac{\partial E}{\partial t} \quad (36)$$

The equation (36), has Ampere as the unit on both the sides, so it may be taken as some sort of current, called Maxwell's displacement current.

On rearranging, equation (36) we have;

$$\frac{\partial q}{\partial t} = A\,\frac{\partial \epsilon_0 E}{\partial t} = A\frac{\partial D}{\partial t} = I_d \quad \text{, (where } D = \epsilon_0 E \text{ is electric displacement vector)} \quad (37)$$

Here, I_d may be called as Maxwell's displacement current.

The equation (37) is the expression for Maxwell's Displacement current. As $D = \epsilon E$ is known as electric displacement vector and this current is proportional to time rate of electric displacement vector, that's why it is called Maxwell's displacement current.

For the consistency of Ampere's circuital law for direct as well as AC circuit, Maxwell modified, the equation (34) as;

$$\oint B.\,dl = \mu_0 (I_C + I_d) \quad (38)$$

Or

$$\oint H.\,dl = (I_C + I_d)$$

Or

$$\oint H.\,dl = I_C + I_d = \oint (J_C + J_d).\,ds \quad (39)$$

The equation (38) or (39), is called modified Ampere's circuital law which is consistent for DC and AC circuits.

*Maxwell's displacement current is different from the conventional current as there is no flow of charge in this case. Now the question that arises is why this is called a current? The explanation to this is that the origin of any kind of magnetic field is current. So, if there is an alternating magnetic field in the free space between the plates of the capacitor, it means that there should be some sort of current. Actually, the flow of charge per unit time is not the proper definition of the current. This can be used to measure the amount of the conduction current but is not the basic definition of current. Whenever charge flow in the conductors, there is a magnetic field around the conductor which means, even in the case of the conduction current, magnetic field always persist. So, existence of alternating magnetic field in the free space between the two plates of conductor, there is always a Maxwell displacement current, whether it is a capacitor or an electric dipole. *

2.15 Maxwell's Equations of Electromagnetism – Foundation of Electromagnetic Waves

Four equations of electrostatics, magnetism, electromagnetic induction with modified Ampere's circuital law, are termed as the Maxwell's equations of electromagnetism.

(i) $\oint E.ds = \frac{q}{\epsilon_0}$ Gauss's Law of Electrostatics

(ii) $\oint B.ds = 0$ Gauss's Law of Magnetostatics

(iii $\xi = \oint E.dl = -\frac{\partial \phi_B}{\partial t}$ Faraday's Law of electromagnetic Induction

(iv) $\oint H.dl = I + I_d = \oint (J + J_d).ds$ Modified Ampere's Circuital Law

These equations are called Maxwell's equations of electromagnetic field in integral from.
Maxwell performed rigorous mathematical operations on these equations and converted these equations into differential form. This helped in interpreting the physical significance of the Maxwell's equations in a very different and an exhaustive manner. He reached on the conclusion that these equations cannot be treated as separated from each other but are very well connected to each other. In the combined form, the physical significance of these equations also gives the foundation of other important physical quantity called as electromagnetic waves.

To convert these equations in differential form, a mathematical operator called Del (∇) is very much needed. Without this operator it is impossible to convert integral form into differential form. So let us have a brief description and discussion about this operator after which it will be very easy to convert the integral form of four equations of electromagnetism into differential form. The physical interpretation of the equations will reveal the hidden concepts of electromagnetic waves in it and finally the foundation for the fundamental equations of electromagnetic waves could be understood.

2.16 Del Operator

The operator, del (or nabla) ' ∇ ' is a mathematical operator, defined as (in rectangular coordinates)

$$\nabla = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)$$

Note: Sometimes the 'del' operator is represented with a vector sign, but it should be kept in mind that 'del' is not a vector quantity; it is an operator which operates in a way similar to a vector.

Some of the operations of $\nabla = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)$ are defined as follows: -

2.16.1 Scalar Multiplication of ∇: Gradient of a Vector

The gradient of a scalar point function f(x, y, z) , is represented by ∇f or $\text{grad}(f)$ and can be obtained by operating the del operator '∇' on the function f. Thus, the gradient of f is given by,

$$\nabla f = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right) f = \left(\frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z}\right) \quad (9)$$

Therefore, gradient of a scalar function is a quantity whose x, y and z components are respectively the partial derivatives of f with respect to x, y and z.

2.16.2 Physical Interpretation of Gradient: If f(x, y, z) is a scalar function at a point (x, y, z)then on changing the x, y and z coordinates by dx, dy, dz the differential change in the scalar function 'f' can be written as,

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \nabla f . d\vec{r}$$

Where, $d\vec{r} = dx\hat{x} + dy\hat{y} + dz\hat{z}$

So, we can write,

$$df = |\nabla f||d\vec{r}|\cos\theta$$

Where θ is the angle between $d\vec{r}$ and ∇f.

From the above equations, we can see that for a fixed value of dr (magnitude of $d\vec{r}$), df will be maximum when $\cos\theta = 1$, i.e., when $d\vec{r}$ is in the direction of ∇f. Therefore, df is maximum in the direction ∇f, so we can say that, the gradient ∇f of a scalar function 'f', points in the direction of maximum increase of the function. Further if $(df)_{max}$ is the maximum change in the function for a fixed value of dr, we can write,

$$(df)_{max} = |\nabla f||d\vec{r}| = |\nabla f|dr$$

Thus, the magnitude of gradient of f is given by,

$$|\nabla f| = \frac{(df)_{max}}{dr} \quad (40)$$

The gradient of a scalar field thus is a vector field which points in the direction of the greatest increase of the scalar field, and whose magnitude is equal to the greatest rate of change of the function with distance.

The gradient can also be used to measure how a scalar field changes in other direction, rather than just the direction of greatest change, by taking its dot product with the unit vector along that direction. Another important property of gradient of a function 'f' is that it is normal to the surface over which f is constant.

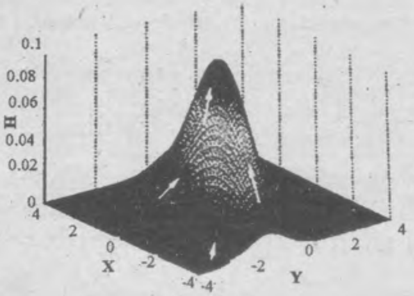


Fig. 2.6

Illustration: Consider that the temperature in a room at different points (x, y, z) of the room is given by (x, y, z). The gradient at each point in the room will show the direction in which the temperature changes most quickly. The magnitude of the gradient determines that how fast the temperature changes in that direction. Further consider a hill whose height at a point (x, y) is H(x, y). Then, the gradient of H at a point is a vector pointing in the direction of the steepest slope at that point (as shown in Fig. 2.6). The steepness of the slope at that point is given by the magnitude of the gradient of H.

2.16.3 Scalar Product of ∇ : Divergence

The divergence of a vector point function $\vec{f}(x,y,z)= f_x\hat{x} + f_y\hat{y} + f_z\hat{z}$ {where f_x, f_y and f_z are respectively the x, y and z components of the vector field $\vec{f}$ at a point having coordinates x, y, z is represented by div($\vec{f}$) or $\nabla.\vec{f}$ and can be obtained by operating the del operator (∇) on the vector field through dot product. Thus, the divergence of the vector field $\vec{f}$ is given by,

$$\nabla.\vec{f} = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right).\vec{f} = \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$$

The divergence of a vector field thus is a scalar quantity.

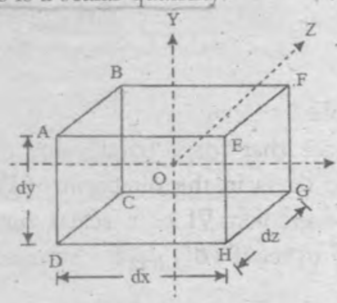


Fig. 2.7

2.16.4 Physical Interpretation of Divergence

Consider an infinitesimal volume with sides dx, dy and dz as shown in Fig. 2.7. If a vector field is $\vec{f}(x,y,z) = f_x\hat{x} + f_y\hat{y} + f_z\hat{z}$ in the middle of the volume at the point 'O', having coordinates (x, y, z), then the x-component of $\vec{f}$ at the middle of the face ABCD, can be of x-component of $\vec{f}$, same over the face thus, the flux of the field $\vec{f}$ through the face ABCD can be written as $\left(f_x - \frac{1}{2}\frac{\partial f_x}{\partial x}dx\right)dxdz$ {the y and z component of $\vec{f}$ do not contribute anything to the flux since these components are perpendicular to the normal to the surface }. Similarly, the x-component of $\vec{f}$ at the face EFGH can be written as $\vec{f}+\frac{1}{2}\frac{\partial f_x}{\partial x}dx$ and the flux through the face EFGH can be written as $\left(\vec{f} + \frac{1}{2}\frac{\partial f_x}{\partial x}dx\right)dydz$. Therefore, the net amount of flux of the field that is diverging from the face EFGH, can be written as , $\left(\vec{f} + \frac{1}{2}\frac{\partial f_x}{\partial x}dx\right)dydz - \left(f_x - \frac{1}{2}\frac{\partial f_x}{\partial x}dx\right)dxdz = \frac{\partial f_x}{\partial x}dxdydz$. Similarly, the flux diverging from the faces ABFE and BCGF, can be written as $\frac{\partial f_y}{\partial y}dxdydz$ and $\frac{\partial f_z}{\partial z}dxdydz$, respectively. Thus, the total flux diverging from the infinitesimal volume is given by, $\left(\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}\right)dxdydz$. The amount of flux diverging through an infinitesimal volume per unit of its volume is given by $\frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} + \frac{\partial f_z}{\partial z}$, which, as we know, is equal to the divergence of the field. We can now define the divergence of a vector field as the net amount of flux of the field diverging through an infinitesimal volume per unit of its volume or we can say that the divergence basically represents the flux generation per unit volume at each point of the field. If 'S' is the surface that bounds a volume 'V' we can alternatively write the divergence of a vector field $\vec{f}$ as,

$$\nabla.\vec{f} = V \rightarrow 0 \frac{\oint_s \vec{f}.d\vec{S}}{V} \quad (41)$$

The divergence of a vector field at a point tells that how much the vector field diverges or spreads out from that point. The point from where the field lines diverge, can be called a source of the field while the point where the field lines converge can be called a sink of the field, so a point of positive divergence is a source and a point of negative divergence is a sink. If the divergence of a vector field is non-zero, then there must be a source or sink of the field.

2.16.5 Solenoidal Vector Fields

The vector fields that have zero divergence everywhere are called solenoidal or divergence-less vector fields.

2.16.6 Gauss' Divergence Theorem

From the above discussion, we have clearly understood that the amount of flux diverging through an infinitesimal volume per unit of its volume is also called divergence of the field and if 'S' is the surface that bounds a volume 'V' we can also write the divergence of a vector field $\vec{A}$ as,

$$\nabla.\vec{A} = V \rightarrow 0 \frac{\oint_s \vec{A}.d\vec{S}}{V}$$

Thus, we can write,

$$\oint_s \vec{A}.\,d\vec{S} = \int_V \left(\nabla.\vec{A}\right)dV \quad (42)$$

In other words, the total inward or outward flux of a vector point function through a closed surface is equal to the divergence of the vector point function, from the volume enclosed by that surface. The equation (42) is known as Gauss' divergence theorem of the vector fields. In simple way it is used to convert any surface integral into volume integral enclosed by the surface.

2.16.7 Vector Product of ∇: Curl

The curl of a vector point function $\vec{f}(x, y, z)=f_x\hat{x} + f_y\hat{y} + f_z\hat{z}$ (where f_x, f_y and f_z respectively are the x, y and z components of the vector field $\vec{f}$ at a point having coordinates x,y,z) is represented by $\nabla \times \vec{f}$ and can be obtained by operating the del operator (∇) on the vector field through cross product. Thus, the curl of $\vec{f}$ is given by,

$$\nabla \times \vec{f} = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right) \times \vec{f}$$

$$= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_x & f_y & f_z \end{vmatrix} = \hat{x}\left(\frac{\partial f_z}{\partial y} - \frac{\partial f_y}{dz}\right) + \hat{y}\left(\frac{\partial f_x}{\partial z} - \frac{\partial f_z}{dx}\right) + \hat{z}\left(\frac{\partial f_y}{\partial x} - \frac{\partial f_x}{dy}\right) \quad (43)$$

Therefore, the curl of a vector field is a vector quantity.

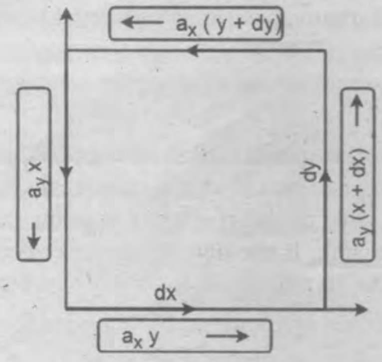


Fig. 2.8

Fig. 2.9

2.16.8 Physical Interpretation of Curl

To see, what the curl of a vector means, consider a rectangular element of length dx and width dy in the region of a vector field $\vec{a}$ as shown in Fig. 2.8. We can write the x-components of the field $\vec{a}$ at the bottom and at the top of the element as,

$a_x(y)$ and $a_x(y+dy) = a_x(y) + \frac{\partial a_x}{\partial y}dy$.

Similarly, we can write, the y-components of the field $\vec{a}$ at the left and right side of the element as,

$a_y(x)$ and $a_y(x+dx) = a_y(x) + \frac{\partial a_y}{\partial x}dx$.

Now working round the clockwise sense, the circulation of the vector around the element can be written as $a_x(y)dx + a_y(x+dx) - a_x(y+dy)dx - a_y(x)dy$, where the minus signs in the last two terms arise because there the path is opposite to the direction of the field.

The circulation of the field $\vec{a}$ around the element can be written as, $a_x(y)dx + \left\{a_y(x) + \frac{\partial a_y}{\partial x}dx\right\}dy - \left\{a_x(y) + \frac{\partial a_x}{\partial y}dy\right\}dx - a_y(x)dy$

Or, $\frac{\partial a_y}{\partial x}dxdy - \frac{\partial a_x}{\partial y}dydx = \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right)dxdy = (\nabla \times \vec{a}).d\vec{S}$ (Where $d\vec{S} = dx\,dy\,\hat{z}$)

The curl of a vector field is thus a measure of the circulation of the vector field per unit area, i.e. also called vertically of the field.

Now the circulation of the vector field around any closed curve can be written as

$\oint_l \vec{a}.d\vec{l} = \oint_l dC = \int_S(\nabla \times \vec{a}).d\vec{S}$

If the integral of a vector field around a closed loop is not zero, then it implies that there is some circulation of the vector field around the loop i.e., a non–zero curl implies that there is a circulation of the vector field. However, if the curl of the vector is zero everywhere, then there cannot be any circulation of the vector field, anywhere in space. Hence, the name 'curl' is given for $\nabla \times \vec{a}$.

The curl of a vector field tells us about the circulation of rotation per unit area the field has at any point.

- The magnitude of the curl tells us how much rotation there is and its direction tells us, by the right – hand rule (four fingers of the right hand are curled in the direction of the vector field, then the thumb points in the direction of the rotation) that about which axis the field is rotating. That is why; curl of a vector field is also called rot (short for rotor).

Illustration: Consider a vector a vector field $\vec{f} = y\hat{x} - x\hat{y}$. This vector field is shown in Fig. 2.9.

From Figure, it can be seen that $\vec{f}$ is circulating around the point 'O'. Using the right – hand rule, we expect the curl to be into the page or in the negative z-direction.

The curl of $\vec{f}$ is given by

$$\nabla \times \vec{f} = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right) \times \hat{f}$$

$$= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & x & 0 \end{vmatrix} = \hat{z}\left(-\frac{\partial x}{\partial x} - \frac{\partial y}{\partial y}\right) = -2\hat{z}$$

It, indeed is in the negative z-direction, as expected. In this case, the curl is actually a constant, irrespective of position. The amount of rotation in the above vector field is the same at any point (x, y).

2.16.9 Irrotational Vector Fields: The vector fields which have zero curl everywhere are called irrotational or curl-less vector fields.

2.16.10 Stoke's Theorem (line to surface integral)

For any vector field $\vec{A}$, we can write as,

$$\int_S (\nabla \times \vec{A}).d\vec{S} = \oint_l \vec{A}.d\vec{l} \quad (44)$$

This equation is called Stokes' theorem.

2.17 Equation of Continuity – Conservation of Charge

The flow of conduction current in a circuit can be written as

$$I = -\frac{\partial q}{\partial t} = -\frac{\partial}{\partial t}\int \rho\, dv \quad (45)$$

volume charge density

We also have;

$$I = \int J.ds \quad (46)$$

From (45) and (46), we have;

$$\int J.ds = -\frac{\partial}{\partial t}\int \rho\, dv$$

$$\int J.ds = \int \frac{\partial \rho}{\partial t} dv$$

Using Gauss's Div. Theorem

$$\int \nabla.J\, dv = -\int \frac{\partial \rho}{\partial t} dv$$

$$\int (\nabla.J + \frac{\partial \rho}{\partial t}) dv = 0$$

As the integral is arbitrary, so the integrand vanishes to zero. i.e.

$$\boxed{\nabla.J + \frac{\partial \rho}{\partial t} = 0} \quad (47)$$

This equation is called the equation of continuity.

2.17.1 Physical Interpretation of Equation of Continuity

The equation (47) shows that if the divergence of conduction current density is zero, i.e., $\nabla.J = 0$ then volume charge density ρ is constant; it means charge is static or charge is conserved. If time rate of change of charge density is zero, then the conduction current density will not be originated from any source. It will flow in a loop like magnetic field lines of force. It means the equation of continuity tells whether the circuit is maintained by AC or DC.

2.18 Maxwell's Equations in Integral Form

The set of four equations of electromagnetism are known as Maxwell's equations in integral form.

(i) $\oint \mathbf{E}.\,d\mathbf{S} = \frac{q}{\epsilon_0}$ Gauss's Law of Electrostatics

(ii) $\oint \mathbf{B}.\,d\mathbf{S} = 0$ Gauss's Law of Magnetostatics

(iii) $\oint \mathbf{E}.\,d\mathbf{l} = -\frac{\partial \phi_B}{\partial t}$ Faraday's Law of EM Induction

(iv) $\oint \mathbf{H}.\,d\mathbf{l} = I + I_d = \oint (J + J_d).\,d\mathbf{S}$ Modified Ampere's Circuital Law

2.18.1 Physical Significance

The physical significance of these is given as;

(i) This equation is Gauss's law in electrostatics which states that, the total outward electric flux over any closed surface is equal to total charge enclosed within volume surrounded by the surface.

(ii) The total outward flux of magnetic flux B through any closed surface 'S' is equal to zero i.e., monopole does not exist.

(iii) If an electric circuit is placed in a magnetic field and the magnetic flux close to a circuit changes, an electromotive force (e.m.f) is induced in the circuit. The magnitude of which is proportional to the rate of change of flux and the direction of the induced e.m.f is given by Lenz's law which states that "the direction of the induced e.m.f. is such that the magnetic flux associated with the current generated by it opposes the original change of flux causing e.m.f".

(iv)This equation has been derived from ampere's law in circuital form for a magnetic field accompanying an electric current. This law states that the line integral of magnetic field around a closed path is equal to the total current crossing any surface bounded by the line integral path.

Most of these equations are more or less as it is, as given by several researchers and does not lead to any specific correlations between them. But when Maxwell converted these equations into differential form, a strong inter-connectivity of these equations come into picture. He also showed that these equations, not only are coupled to each other but are the basic equations of electromagnetic field theory. These equations can then be used further to derive the equations of electromagnetic waves.

2.18.2 Conversion of Maxwell's equations - Integral Form into Differential Form

1. **Conversion of Gauss law of electrostatics**

Here we have, $\oint \mathbf{E}.\,d\mathbf{S} = \frac{q}{\epsilon_0}$

If ρ is the charge density, then $\rho\, dV$ is the charge in volume element dV which contributes $\frac{1}{\epsilon_0}\rho\, dV$ to the surface integral. If the surface 's' encloses a volume V then

$$\oint \mathbf{E}.\,d\mathbf{S} = \frac{1}{\epsilon_0}\oint \rho\, dV$$

But if the surface is bounded volume V of a dielectric, then the total charge must include both the free and polarisation charges. Thus, the total charge density at a point in a small volume dV, should be $(\rho + \rho')$ where ρ' is polarisation charge density and ρ is free charge density. Thus, the above equation can be expressed as

$$\oint \mathbf{E}.\,d\mathbf{S} = \frac{1}{\epsilon_0}\oint (\rho + \rho')\, dV$$

If P is polarisation i.e., electric dipole moment per unit volume then div P or $\nabla \cdot P$ is the amount of polarised charge in a unit volume. As the polarised charge is reverse in nature with respect to real charge, thus

$$\rho' = -\nabla \cdot P = -\,\mathrm{div}\,P$$

$$\oint \mathbf{E}.\,d\mathbf{S} = \frac{1}{\epsilon_0}\oint (\rho - \mathrm{div}\,P)\,dV$$

$$\oint \mathbf{E}.\,d\mathbf{S} = \frac{1}{\epsilon_0}\oint \rho\,dV - \oint \mathrm{div}\,P\,dV$$

Applying Gauss Divergence Theorem to change surface integral to volume integral, we get

$$\oint \epsilon_0\mathbf{E}.\,d\mathbf{S} = \oint \mathrm{div}\,(\epsilon_0\mathbf{E})dV = \oint \rho dV - \oint \mathrm{div}\,P\,dV$$

$$\oint \mathrm{div}(\epsilon_0\mathbf{E} + \mathbf{P})dV = \oint \rho\,dV$$

The quantity, $\epsilon_0\mathbf{E} + \mathbf{P} = \mathbf{D}$, called electric displacement vector,

$$\oint \mathrm{div}\,\mathbf{D}\,dV = \oint \rho\,dV$$

$$\oint \mathrm{div}(\mathbf{D} - \rho)dV = 0$$

Since the equation is true for arbitrary volume, the integrand must vanish,

Thus

$$\mathrm{div}(\mathbf{D} - \rho) = 0$$

$$\mathrm{div}\,\mathbf{D} = \rho$$

$$\nabla.\mathbf{D} = \rho \tag{48}$$

Now we have, $\mathbf{D} = \epsilon\mathbf{E}$ where ϵ is the permittivity of the dielectric medium.

In free space, $\mathbf{D} = \epsilon_0\mathbf{E}$, ϵ_0 is the permittivity of free space.

Thus,

$$\boxed{\nabla.\mathbf{E} = \frac{\rho}{\epsilon_0}} \tag{49}$$

Equations (48) or (49) is first equation in differential form.

2. Conversion of Gauss law of Magnetostatics

Since the magnetic lines of force are either closed or go off to infinity, the number of magnetic lines of force entering any arbitrary close surface is exactly the same as leaving it. This means the total outward flux of magnetic induction B through any closed surface 's' is equal to zero i.e.

$$\oint \mathbf{B}.\,d\mathbf{S} = 0$$

Transforming the surface integral into volume integral, using Gauss divergence theorem, we get

$$\oint \nabla.\mathbf{B}\,dV = 0$$

The integrand should vanish for the surface boundary as the volume is arbitrary i.e.

$$\mathrm{div}\,\mathbf{B} = 0$$

or

$$\boxed{\nabla.\mathbf{B} = 0} \tag{50}$$

This is differential form of second equation.

3. Conversion of Faradays law of electromagnetic induction and Lenz law

We know according to Faraday's law of electromagnetic induction, the induced e.m.f is given by;

$\xi = -\frac{\partial\phi_B}{dt}$,

i.e., time rate of change of magnetic flux produces an induced e.m.f. and it is produced to oppose this change. Also, the induced e.m.f. is equal to the work done per unit charge and can be written as the line integral of the induced electric field E around the circuit i.e.

$e = \oint \mathbf{E}.\,\mathbf{dl}$

And the rate of change of magnetic flux through the circuit is equal to

$$\frac{\partial \phi_B}{dt} = \oint \frac{\partial (\mathbf{B}.\,\mathbf{dS})}{\partial t}$$

Where the integral is taken over any area 's' bounded by the circuit. Since the surface ds does not change its shape or position with time, we can write the above equation as:

$$\oint \mathbf{E}.\,\mathbf{dl} = -\oint \frac{\partial \mathbf{B}}{\partial t}.\,\mathbf{dS}$$

The total time derivative has been changed to partial derivative as we are only concerned with the changes in the field **B** with time at a fixed position of the elemental area. it signifies that "the electromotive force around a closed path is equal to the time derivative of the magnetic displacement through any surface bounded by the path.

Using Stokes theorem, the line integral can be transformed into the surface integral i.e.

$\oint \mathbf{E}.\,\mathbf{dl} = \oint \text{curl}\,\mathbf{E}.\,\mathbf{dS} = -\oint \frac{\partial \mathbf{B}}{\partial t}.\,\mathbf{dS}$

Or $\oint (\text{curl}\,\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t}).\,\mathbf{dS} = 0$

This equation must hold for any arbitrary surface in the field, thus the integrand should vanish i.e.

$\text{curl}\,\mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0$

Or

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}} \qquad (51)$$

This is called differential form of Maxwell third equation.

4. Conversion of Ampere's Circuital Law

Ampere's law states that the line integral of magnetic field intensity around a closed path is equal to the total current crossing any surface bounded by the line integral path i.e.

$$\oint \mathbf{H}.\,\mathbf{dl} = \mathrm{I} = \oint \mathbf{J}.\,\mathbf{dS}$$

Now changing the line integral into the surface integral by the use of Stoke's theorem,

$\oint \mathbf{H}.\,\mathbf{dl} = \oint \text{curl}\,\mathbf{H}.\,\mathrm{dS} = \oint \mathbf{J}.\,\mathbf{dS}$ Or

$\oint (\nabla \times \mathbf{H} - \mathbf{J}).\,\mathbf{dS} = 0$

For an arbitrary surface, the integrand should vanish, thus

$\nabla \times \mathbf{H} - \mathbf{J} = 0$, Or

$\boxed{\nabla \times \mathbf{H} = \mathbf{J}}$, Or

$$\text{curl}\,\mathbf{H} = \mathbf{J} \qquad (52)$$

This is the differential form of Ampere's law for steady current only.

2.18.2.1 Consistency of Ampere's Law for Time Varying Fields on the basis of Maxwell's equations

This is another way to check the consistency of Ampere's law for DC and AC fields. As we can see from the equation of continuity, the divergence of current density can reveal whether the field is DC or AC. So, taking the divergence of equation (23), we have;

$\nabla.(\nabla \times \mathbf{H}) = \nabla.\mathbf{J}$

$(\nabla.\nabla)\,\mathbf{H} - \nabla(\nabla.\mathbf{H}) = \nabla.\mathbf{J},$ Or

$0 = \nabla.\mathbf{J}$

It means the divergence of current density is zero.

According to the equation of continuity;

$$\nabla.\mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{53}$$

If $\nabla.\mathbf{J} = 0$, then ρ is constant i.e., volume charge density is not function of time. This shows that the Ampere's circuital is true only for the circuits in which charge does not change with time i.e., it is true for direct fields only.

To make it consistent for a circuit having connected a capacitor or an electric dipole maintained by alternating electric filed, Maxwell added some quantity in the right-hand side of the equation (52) let it be $\mathbf{J}'$. So, the Ampere's equation is modified as;

$$\text{curl}\,\mathbf{H} = \mathbf{J} + \mathbf{J}' \tag{54}$$

As, the addition is in current density so $\mathbf{J}'$is also some kind of current density. Now the divergence of equation (54) gives as;

$\nabla.(\text{curl}\,\mathbf{H}) = \nabla.(\mathbf{J} + \mathbf{J}')$

Or

$$\nabla.\mathbf{J} + \nabla.\mathbf{J}' = 0 \tag{55}$$

Using equation (53) and (55);

$$\nabla.\mathbf{J}' = \frac{\partial \rho}{\partial t} \tag{56}$$

From (48) and (56);

$\nabla.\mathbf{J}' = \frac{\partial(\nabla.\mathbf{D})}{\partial t}$

Or

$\nabla.\mathbf{J}' = \nabla.\frac{\partial \mathbf{D}}{\partial t}$

Or

$\mathbf{J}' = \frac{\partial \mathbf{D}}{\partial t}$ or

$$\mathrm{I}' = \mathrm{A}\frac{\partial \mathbf{D}}{\partial t} \tag{57}$$

As the current given by the equation (57), is directly proportional to the time rate of change of electric displacement vector $\mathbf{D}$, this may be called as displacement current or Maxwell's displacement current.

In space, the magnitude of Maxwell's displacement current can be calculated as;

$\mathrm{I_d} = \mathrm{A}\epsilon_0 \frac{\partial \mathrm{E}}{\partial t}$

Where A is the area, ϵ_0 is the permittivity of free space where the electric field is changing with time.

So, the differential form of modified Ampere's law can be written as the Maxwell's fourth equation of electromagnetic wave in differential form. That is;

$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{J}' \tag{58}$$

Here, it is very important to note that the equation (58) is true for a DC circuit as well as AC circuit having a capacitor like component or an electric dipole operated by an alternating electric field (Fig.2.5). In the portion of the circuit containing conducting wires only, i.e., in the portion of circuit containing conduction current, only the conduction current density comes into picture and the displacement current density will be zero. But in the portion of the circuit where there is free space, for example the space between the plates of a capacitor or space between the two poles of an electric dipole operated by alternating electric field, only the displacement current density will come into picture and at that place the conduction current density will be zero. From this, it may be concluded that in a circuit shown in Fig. 2.5, the magnitude of conduction current and Maxwell current should be equal as it may be considered that the current in whole of the circuit remains the same. In the wires, it is conduction current I_C but in the free space it is Maxwell displacement current I_d.

2.18.3 Physical Significance of Maxwell Equations: Differential Form

Maxwell equations in differential form are given as;

$$\left.\begin{aligned}
&\nabla.\mathbf{E} = \frac{\rho}{\epsilon_0} \text{ or } \nabla.\mathbf{D} = \rho && \text{(i)}\\
&\nabla.\mathbf{B} = 0 && \text{(ii)}\\
&\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = \mu_0 \frac{\partial \mathbf{H}}{\partial t} && \text{(iii)}\\
&\nabla \times \mathbf{H} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} && \text{(iv)}
\end{aligned}\right\} \quad (59)$$

Physical Significance

$\Delta \cdot D$ ρ

1. The divergence of electric displacement vector gives the volume charge density. If it is positive than the position of charge is the origin of source of the electric field, if it is negative, the electric field is sinking at the position of charge and if it is zero then the electric field entering from one side is the same as coming out from the other side. But if the value of charge divergence is oscillating between a maximum and minimum value then the charge density is changing with time i.e., there is oscillating electric field at that point.
2. The divergence of magnetic field at any point is always zero. This verifies the concept that monopole does not exist.
3. This equation shows that curling effect of electric field give us the changing magnetic field with time and also perpendicular to the electric field. This also shows that space varying electric field gives time varying magnetic field.
4. This equation is one of the most important contributions of Maxwell. If the Ampere's law was not modified by Maxwell, the equation 59(iv) shows that space varying magnetic field gives conduction current density and there is no time varying electric field. After the introduction of displacement current, it is possible that space varying magnetic field can generate time varying electric field in the free space. And that is the basic concept of electromagnetic waves.

The physical significance of the equation 59(iii) and 59(iv) is the foundation for the development of the theory of electromagnetic waves. As 59(iii) and 59(iv) equations show vice a versa effects of varying electric and magnetic fields i.e., if magnetic field vary at any point in space, it propagates varying electric field in the space around it and if electric field vary with time at any point, the varying magnetic field propagates in the space around it. It means that propagation of varying electric and magnetic field is due to the time variation of magnetic and electric field respectively. As the propagation of electromagnetic wave in any medium is the propagation of varying electric and magnetic field in the medium, so we can conclude that these four equations of Maxwell are the fundamental equations which can be used for further derivation of equations of electromagnetic waves in different media. Now let us derive the equations of plane electromagnetic waves in various media.

2.19 Work done by the electromagnetic field

The propagation of electromagnetic waves i.e., oscillating electric and magnetic fields, in free space or in any medium is possible on account of the energy carried by with the wave. This is generated due to the work done by the electromagnetic field.

2.19.1 Energy of Electromagnetic Field

Like all other types of waves, electromagnetic waves also transport energy as they travel from one place to other. As, the energy, in an electromagnetic system is due to the oscillating electric and magnetic fields. So, the total energy associated with electromagnetic wave will be;

$$u = \frac{1}{2}\epsilon_0 E^2 + \frac{1}{2\mu_0}B^2$$

We have, $B = \frac{E}{C}$, so

$$u = \frac{1}{2}\epsilon_0 E^2 + \frac{1}{2\mu_0}\frac{E^2}{C^2},$$

Using $C = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$, $u = \frac{1}{2}\epsilon_0 E^2 + \frac{1}{2}\epsilon_0 E^2 = \epsilon_0 E^2$

This also proves that the electric field energy is equal to magnetic field energy in an electromagnetic wave.

2.20 Propagation of Electromagnetic Field Energy – Poynting Vector

Now we have to find out, how and in which direction the electromagnetic field energy propagates. Let a plane electromagnetic wave of cross-sectional area ds, travelling with velocity C. The volume, crossing in unit time will be C.d**S**. So, the energy density passing with this region per second is;

$$d\mathbf{U} = u\,C\,d\mathbf{S} = \epsilon_0 E^2 C\,d\mathbf{S}$$

So, the energy flowing per time per unit cross-sectional area will be;

$$\mathbf{S} = \frac{dU}{dS} = \epsilon_0 E^2 C = uC \tag{60}$$

$$= \epsilon_0 E^2 . \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

$$= E^2 \sqrt{\frac{\epsilon_0}{\mu_0}}$$

$$= E^2 \frac{1}{\mu_0}\sqrt{\mu_0 \epsilon_0}$$

$$= \frac{EE}{C}\frac{1}{\mu_0} \qquad \text{(since } \frac{E}{B} = C\text{)}$$

So, $\boxed{S = \frac{1}{\mu_0}EB}$

As we know in an electromagnetic wave, electric and magnetic field vectors are perpendicular to each other, so the above expression, in vector form can be written as;

$$\mathbf{S} = \frac{1}{\mu_0}(\mathbf{E} \times \mathbf{B}) \tag{61}$$

Here **S** is a vector form of flow of electromagnetic energy per unit cross-sectional area per unit time, which is perpendicular to both the electric and magnetic field vectors, called Poynting vector, also can be written as;

$$\mathbf{S} = \mathbf{E} \times \mathbf{H} \tag{62}$$

$H = \frac{B}{\mu_0}$

2.21 Electromagnetic Energy Theorem - Poynting Theorem

Poynting theorem in electromagnetism is equivalent to mass - energy theorem in mechanics. It states that "rate of decrease of the energy in the electromagnetic fields in any volume minus flow of energy per second through the surface of that volume is equal to the rate of work done on the charges by the electromagnetic force or power transferred in the space". It can be written as

$$-\frac{\partial}{\partial t}\oint_V\left(\frac{1}{2}\frac{B^2}{\mu_0}+\frac{1}{2}\epsilon_0 E^2\right)dV-\frac{1}{\mu_0}\oint(\mathbf{E}\times\mathbf{B}).\,d\mathbf{S}=\oint_V(\mathbf{E}.\,\mathbf{J})dV$$

$$\text{Or} -\frac{\partial}{\partial t}\oint_V\left(\frac{1}{2}\mathbf{E}.\,\mathbf{D}+\frac{1}{2}\mathbf{H}.\,\mathbf{B}\right)-\oint(\mathbf{E}\times\mathbf{H}).\,d\mathbf{S}=\oint_V(\mathbf{E}.\,\mathbf{J})dV \tag{63}$$

Where

(i) $-\frac{\partial}{\partial t}\oint_V\{\frac{1}{2}\frac{B^2}{\mu_0}+\frac{1}{2}\epsilon_0 E^2\}dV$ = Rate of decrease of electromagnetic energy in the electromagnetic fields in the volume V

(ii) $-\frac{1}{\mu_0}\oint(\mathbf{E}\times\mathbf{B}).\,d\mathbf{S}$ = Rate at which the energy is propagated by the electromagnetic fields through the cross-sectional area that bounds the volume V (Energy flow / Area)

(iii) $\oint_V(\mathbf{E}.\,\mathbf{J})dV$ = Power transferred into the free space or rate of work done by the electromagnetic fields in the volume V (Power loss)

2.21.1 Derivation of Poynting Theorem

Writing Maxwell's equations in differential form

$$\left.\begin{aligned}
&\nabla.\,\mathbf{D}=\rho \quad \text{or} \quad \text{div}\,\mathbf{D}=\rho && \text{(i)}\\
&\nabla.\,\mathbf{B}=0 \quad \text{or} \quad \text{div}\,\mathbf{B}=0 && \text{(ii)}\\
&\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} \quad \text{or} \quad \text{curl}\,\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} && \text{(iii)}\\
&\nabla\times\mathbf{H}=\mathbf{J}+\frac{\partial\mathbf{D}}{\partial t} \quad \text{or} \quad \text{curl}\,\mathbf{H}=\mathbf{J}+\frac{\partial\mathbf{D}}{\partial t} && \text{(iv)}
\end{aligned}\right\} \tag{64}$$

Taking scalar product of equation {64(iii)} with **H** and equation {64(iv)} with **E**, we get;

$$\mathbf{H}.\,(\nabla\times\mathbf{E})=-\mathbf{H}.\frac{\partial\mathbf{B}}{\partial t} \tag{65}$$

$$\mathbf{E}.\,(\nabla\times\mathbf{H})=\mathbf{E}.\,\mathbf{J}+\mathbf{E}.\frac{\partial\mathbf{D}}{\partial t} \tag{66}$$

Now subtracting equation (66) from equation (65), we get;

$$\mathbf{H}.\,(\nabla\times\mathbf{E})-\mathbf{E}.\,(\nabla\times\mathbf{H})=-\mathbf{H}.\frac{\partial\mathbf{B}}{\partial t}-\mathbf{E}.\frac{\partial\mathbf{D}}{\partial t}-\mathbf{E}.\,\mathbf{J}$$

$$\mathbf{H}.\,(\nabla\times\mathbf{E})-\mathbf{E}.\,(\nabla\times\mathbf{H})=-\left(\mathbf{H}.\frac{\partial\mathbf{B}}{\partial t}+\mathbf{E}.\frac{\partial\mathbf{D}}{\partial t}\right)-\mathbf{E}.\,\mathbf{J} \tag{67}$$

We have the vector identity;

$$\nabla.\,(\mathbf{E}\times\mathbf{H})=\mathbf{H}.\,\text{curl}\,\mathbf{E}-\mathbf{E}.\,\text{curl}\,\mathbf{H}$$

So, the equation (67) takes the form;

$$\nabla.\,(\mathbf{E}\times\mathbf{H})=-\left(\mathbf{H}.\frac{\partial\mathbf{B}}{\partial t}+\mathbf{E}.\frac{\partial\mathbf{D}}{\partial t}\right)-\mathbf{E}.\,\mathbf{J} \tag{68}$$

For a linear medium we can write;

$$\mathbf{B}=\mu\mathbf{H} \text{ and } \mathbf{D}=\epsilon\mathbf{E}$$

So, the equation (68) be written as;

$$\nabla.\,(\mathbf{E}\times\mathbf{H})=-\left[\mathbf{H}.\frac{\partial(\mu\mathbf{H})}{\partial t}+\mathbf{E}.\frac{\partial(\epsilon\mathbf{E})}{\partial t}\right]-\mathbf{E}.\,\mathbf{J} \tag{69}$$

Now rearranging $\mathbf{E}.\frac{\partial(\epsilon\mathbf{E})}{\partial t}$ and $\mathbf{H}.\frac{\partial(\mu\mathbf{H})}{\partial t}$, we have;

$$\mathbf{E}.\frac{\partial}{\partial t}(\epsilon\mathbf{E})=\frac{1}{2}\epsilon\frac{\partial}{\partial t}(E^2)=\frac{\partial}{\partial t}\left(\frac{1}{2}\,\epsilon E^2\right)=\frac{\partial}{\partial t}\left(\frac{1}{2}\,\mathbf{E}.\,\mathbf{D}\right) \tag{70}$$

$$\mathbf{H}.\frac{\partial}{\partial t}(\mu\mathbf{H}) = \frac{1}{2}\mu\frac{\partial}{\partial t}(H^2) = \frac{\partial}{\partial t}\left(\frac{1}{2}\mu H^2\right) = \frac{\partial}{\partial t}\left(\frac{1}{2}\ \mathbf{H}.\mathbf{B}\right) \qquad (71)$$

Using expression (70) and (71), the equation (69) takes the form;

$$\nabla.(\mathbf{E}\times\mathbf{H}) = -\frac{\partial}{\partial t}\left[\frac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B})\right] - \mathbf{J}.\mathbf{E} \qquad (72)$$

On integration, the equation (52) over a volume V bounded by surface S, we have;

$$\int_V \nabla.(\mathbf{E}\times\mathbf{H})\,d\mathbf{V} = -\int_V \left\{\frac{\partial}{\partial t}\left(\frac{1}{2}\mathbf{E}.\mathbf{D}+\frac{1}{2}\mathbf{H}.\mathbf{B}\right)\right\}d\mathbf{V} - \int_V \mathbf{J}.\mathbf{E}\,d\mathbf{V}$$

Using Gauss divergence theorem to change volume integral on L.H.S. of the above equation into surface integral, we get

$$\int_V \mathbf{J}.\mathbf{E}\,d\mathbf{V} = -\frac{\partial}{\partial t}\int_V\left(\frac{1}{2}\mathbf{E}.\mathbf{D}+\frac{1}{2}\mathbf{H}.\mathbf{B}\right)d\mathbf{V} - \oint_s(\mathbf{E}\times\mathbf{H}).d\mathbf{S} \qquad (73)$$

This is the Poynting theorem.

2.21.2 Physical Significance of Poynting Theorem

Let us interpret each of the three terms individually.

(i) Physical Interpretation of $-\oint_s(\mathbf{E}\times\mathbf{H}).d\mathbf{S}$

$\mathbf{E}\times\mathbf{H} = \mathbf{S}$ is a quantity called poynting vector and is defined as the energy flow per unit time per unit area i.e., power flow per unit area and is perpendicular to **E and H** both. So $\oint_s(\mathbf{E}\times\mathbf{H}).d\mathbf{S}$ may be defined as the rate at with the energy of the electromagnetic fields is decreased through the cross-sectional area that bounds the volume V

(ii) Physical Interpretation of $-\frac{\partial}{\partial t}\int_V\left(\frac{1}{2}\mathbf{E}.\mathbf{D}+\frac{1}{2}\mathbf{H}.\mathbf{B}\right)dV$

Here, we have;

$\int_V \frac{1}{2}\mathbf{E}.\mathbf{D}\,d\mathbf{V} = U_e,$ Electrostatic potential energy in volume V

$\int_V \frac{1}{2}\mathbf{H}.\mathbf{B}\,d\mathbf{V} = U_m,$ Magnetic energy in volume V

$U = \int_V \frac{1}{2}(E.D + H.B)dV$, represents some sort of potential energy of electromagnetic field, as it exists due to static fields (electric and magnetic), known as electromagnetic field energy in volume V. A concept such as energy stored in the field itself rather than residing with the particles is a basic concept of electromagnetic theory.

Obviously $\frac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B})$ represents energy density of electromagnetic field i.e.

$$U = \frac{1}{2}(\mathbf{E}.\mathbf{D}+\mathbf{H}.\mathbf{B})$$

Consequently, the term, $-\frac{\partial}{\partial t}\int_V\left(\frac{1}{2}\mathbf{E}.\mathbf{D}+\frac{1}{2}\mathbf{H}.\mathbf{B}\right)d\mathbf{V}$ represents the rate of decrease of stored electromagnetic energy in volume V.

(ii) Physical Interpretation of $\int_V \mathbf{J}.\mathbf{E}\,dV$

To understand the physical significance of this term, let us consider a charged particle q in the charge distribution acted upon by Lorentz force $\mathbf{F} = q(\mathbf{E}+\mathbf{v}\times\mathbf{B})$, displaced by an amount $\mathbf{dl}$. Where $\mathbf{v}$ may be considered as the drift velocity of the charges. So, the work done in this displacement will be given as;

$dW = \mathbf{F}.d\mathbf{l} = q(\mathbf{E}+\mathbf{v}\times\mathbf{B}).\mathbf{dl} = q(\mathbf{E}+\mathbf{v}\times\mathbf{B}).\mathbf{v}dt$

Or $dW = q\mathbf{E}.\mathbf{v}dt = \mathbf{F}.\mathbf{v}dt$

Or $\frac{\partial W}{\partial t} = \mathbf{F.v}$

We can also get the above expression as for an electromagnetic force due to field vectors E and B acting on the charged particle, the magnetic force $\mathbf{q(v \times B)}$ is always perpendicular to velocity. Hence, the magnetic field does no work. Therefore, for a single charge q the rate of doing work by electromagnetic field E and B is;

$$\frac{\partial W}{\partial t} = \mathbf{F.v} = q\mathbf{E.v} \qquad (74)$$

If an electromagnetic field consists of a group of charges moving with different velocities e.g. n_i charge carriers each with charge q_i, moving with velocity v_i (i-1,2,3...); then equation (54) can be written as;

$$\frac{\partial W}{\partial t} = -\Sigma\, n_i q_i \mathbf{v}_i.\mathbf{E}_i \qquad (75)$$

In this case, the total current density $\mathbf{J} = \Sigma\, \mathbf{J}_i = \Sigma\, n_i q_i \mathbf{v}_i$. So, equation (55) becomes

$$\frac{\partial W}{\partial t} = -\Sigma\, \mathbf{J}_i.\mathbf{E}_i = -\mathbf{J}.\mathbf{E} \qquad (76)$$

Therefore, the expression, $\int_V \mathrm{J.E\,dV}$ represents rate of energy transferred into the electromagnetic field through the motion of free charge in volume V.

The physical significance of the equation of Poynting theorem is that the time rate of decrease of electromagnetic energy with a certain volume plus time rate of the energy flowing out through the boundary surface is equal to the power transferred into the electromagnetic field. This is also the statement of conservation of energy in electromagnetism which is known as Poynting theorem.

2.22 Momentum of Electromagnetic Field

As we know, anything containing energy, also have momentum. So, electromagnetic waves also must have momentum. The expression for momentum of electromagnetic waves can be derived with the analogy of momentum of a d' Broglie wave.

We have $P = \frac{h}{\lambda}$, here P is the momentum of the d' Broglie wave, λ is the wavelength and h is Planck's constant.

$$u = h\nu = h\frac{C}{\lambda}$$

Or $u = PC$

$$\text{Or } P = \frac{u}{C} \qquad (77)$$

From equation (60), we have the energy density per unit area per unit time as;

$S = uC$ (here u is the energy density)

$$\text{So } \frac{u}{C} = \frac{S}{C^2} \qquad (78)$$

From (77) and (78), we can have the momentum density of electromagnetic waves as

$$\mathbf{P} = \frac{u}{C} = \frac{S}{C^2} \qquad (79)$$

Or

$$\mathbf{P} = \frac{1}{\mu_0 C^2}(\mathbf{E} \times \mathbf{B}) \qquad \text{as } \mathbf{S} = \frac{1}{\mu_0}(\mathbf{E} \times \mathbf{B}) \qquad (80)$$

Or

$$\boxed{\mathbf{P} = \frac{1}{C^2}(\mathbf{E} \times \mathbf{H})} \qquad (81)$$

Using equation (80) or (81), the momentum transferred by the electromagnetic waves into the medium, can be calculated, if we have electric and magnetic field vectors at that point of the medium.

2.23 Angular Momentum in Electromagnetic Fields

Angular momentum is a physical property of a rotating object. Now, the first question arises that what is the concept behind the angular momentum of electromagnetic waves? The propagation of a beam of an electromagnetic wave is considered due to the helical type rotation of electric as well as magnetic fields. So, if we look through the cross section of a light beam, the propagation of the beam of electromagnetic wave will be seen as, composed of helical as well as spin type motion of light beam. That's why electromagnetic wave should also possess angular momentum.

As according to the fundamental definition of angular momentum, "moment of linear momentum is called angular momentum of a rotating object". So, according to this definition, the angular momentum of beam of electromagnetic waves can be calculated as;

$$\mathbf{L} = \mathbf{r} \times \mathbf{P}$$

Here P is the linear momentum of electromagnetic field/wave and L is the moment of linear momentum about the axis of rotation having r as the distance from axis of rotation, called angular momentum. Now substituting the value of 'P' from (79) into the above equation, we have;

$$\mathbf{L} = \frac{1}{C^2}(\mathbf{r} \times \mathbf{S}) \text{ , or}$$

$$\mathbf{L} = \frac{1}{\mu_0 C^2}(\mathbf{r} \times \mathbf{E} \times \mathbf{B}) \qquad (82)$$

$$\mathbf{L} = \frac{1}{C^2}(\mathbf{r} \times \mathbf{E} \times \mathbf{H}) \qquad (83)$$

Here, it should be noted that angular momentum calculated from (82) or (83), is the angular momentum density. The actual angular momentum will be found out using;

$$\mathbf{L} = \frac{1}{\mu_0 C^2}\int(\mathbf{r} \times \mathbf{E} \times \mathbf{B})\, dv \qquad (84)$$

$$\mathbf{L} = \frac{1}{C^2}\int(\mathbf{r} \times \mathbf{E} \times \mathbf{H})\, dv \qquad (85)$$

Obviously, the angular momentum of electromagnetic field, at any point can be calculated by taking moment of the cross product of electric and magnetic field vectors of an electromagnetic wave at that point. From equations (84) or (85), it can also be seen that angular momentum of an electromagnetic wave is always perpendicular to the direction of propagation of electromagnetic wave.

The propagation of a beam of an electromagnetic wave, as viewed through the cross section of the beam, is composed of two types of motion, one as the helical motion called orbital motion and the other is the spin on its own axis. The beam of electromagnetic waves can be actually considered as rotating around its own axis while propagating in helical path. So, the angular momentum of an electromagnetic wave should be composed of spin angular momentum (SAM) and orbital angular momentum (OAM) both. i.e.

$$\mathbf{L} = \mathbf{L}_{\mathbf{SAM}} + \mathbf{L}_{\mathbf{OAM}}$$

For a well collimated beam, the optical polarization also called circular polarization is exclusively due to Spin Angular Momentum while the Orbital Angular Momentum is related with the spatial field distribution, and in particular with the wave-front helical shape. However, for highly focused or diverging beam or otherwise in general, total angular momentum '**L**', may server the purpose. Spin angular momentum is widely being used in radar applications while orbital angular momentum is being employed in optical fibre transmission.

2.24 Idea of Electromagnetic Waves

Ampère's circuital law was written and well explained for a circuit in which a conduction current is flowing in the circuit. But if a circuit has a region where no conventional current is flowing, the Ampere circuital law will not be valid at that place for example, the Ampere's circuital law cannot be applied throughout the whole circuitry, if a capacitor is connected in a circuit powered by an alternating field

Fig.2.1. During the flow of current inside the circuit, the Ampere's circuital law is applicable throughout the whole circuit except in the space between the plates of the capacitor as there is no conventional current. So, the Ampere's circuital law i.e., $\oint \mathbf{B}.\mathbf{dl} = \mu_0 \mathbf{I} = 0$ and is not valid in the space between the plates of a capacitor. But the circuit is complete otherwise current will not flow in the connecting wires. It means two plates of the capacitor should be connected with each other internally so that that the current may flow in the external circuit. As, the whole circuit contains conduction current except the gap between the plates of the capacitor, Maxwell gave a hypothesis that there must be some kind of current in between the plates of the capacitor. Later, he termed it Maxwell displacement current which was due to the presence of variable electric and magnetic fields between the plates. It means some energy propagates from one plate to the other through the free space between the plates during the charging and discharging of the capacitor over the complete cycle of the alternating field. As this energy is propagating in the form of oscillating electric and magnetic field in the free space between the plates, this may be termed as electromagnetic waves.

2.24.1 Discovery of Electromagnetic Waves

The existence of electromagnetic waves was first investigated by Heinrich Hertz, who succeeded in generating and detecting radio waves. But he could not lay down the fundamental formulations for the electromagnetic waves. Maxwell is well known for his pioneer work in the area of electromagnetic waves. Using fundamental equations of electromagnetism, he not only developed the equations for electromagnetic waves but also proved that light is an electromagnetic wave and is transverse in nature. Today, we all know about complete electromagnetic wave spectrum which we receive from the Sun.

2.25 Electromagnetic Waves in Free Space

Although the four equations written in differential form are sufficient to understand the concept of electromagnetic waves, but there is a need to formulate the equation of electromagnetic waves which resembles the general wave equation so that we can treat and analyse that equation to find various physical parameters of a wave. By doing some mathematical operations on the Maxwell equation in differential form, we can obtain the required equations of electromagnetic waves.

General Maxwell equations in differential form are as follows;

$$\left.\begin{array}{ll} \nabla.\mathbf{E} = \frac{\rho}{\epsilon_0} \text{ or } \nabla.\mathbf{D} = \rho & \text{(i)} \\ \nabla.\mathbf{B} = 0 & \text{(ii)} \\ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} & \text{(iii)} \\ \nabla \times \mathbf{H} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} & \text{(iv)} \end{array}\right\} \quad (86)$$

For free space, the charge density ρ and conduction current density J are both zero. So, the Maxwell's equations reduce to:

$$\left.\begin{array}{ll} \nabla.\mathbf{E} = 0 & \text{(i)} \\ \nabla \times \mathbf{B} = 0 & \text{(ii)} \\ \nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} & \text{(iii)} \\ \nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} & \text{(iv)} \end{array}\right\} \quad (87)$$

Where μ_0 and ϵ_0 is the permeability and permittivity of free space.

Now taking curl of the equations {(87)(iii)}, we have

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu_0 \frac{\partial (\nabla \times \mathbf{H})}{\partial t} \quad (88)$$

Simplifying the vector triple product and putting the value of $\nabla \times \mathbf{H}$ from {(87)(iv)} into (88);

$$\nabla.(\nabla.\mathbf{E}) - \nabla^2\mathbf{E} = \mu_0 \frac{\partial}{\partial t}[\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}]$$

$$\nabla^2\mathbf{E} - \mu_0\epsilon_0 \frac{\partial^2\mathbf{E}}{dt^2} = 0 \qquad \text{(As } \nabla.\mathbf{E} = 0 \text{ for free space)} \qquad (89)$$

Similarly, taking curl of the equations {(87)(iv)}, we have

$$\nabla \times (\nabla \times \mathbf{H}) = -\epsilon_0 \frac{\partial(\nabla \times \mathbf{E})}{\partial t} \qquad (90)$$

Simplifying the vector triple product and putting the value of $\nabla \times \mathbf{E}$ from {(87)(iii)} into (90);

$$\nabla.(\nabla.\mathbf{H}) - \nabla^2\mathbf{H} = -\epsilon_0\mu_0 \frac{\partial}{\partial t}[\frac{\partial \mathbf{H}}{\partial t}]$$

$$\nabla^2\mathbf{H} - \mu_0\epsilon_0 \frac{\partial^2\mathbf{H}}{dt^2} = 0 \qquad (91)$$

The equations (89) and (91) resemble the general wave equation (28), generally written as;

$$\nabla^2\mathbf{u} - \frac{1}{v^2}\frac{\partial^2\mathbf{u}}{dt^2} = 0 \qquad (92)$$

Where $\mathbf{u}$ is a wave variable and $\mathbf{v}$ is the velocity of the wave.

Now comparing the equations (89) and (91) with (92), we can say that **E** or **H** are the variables of the wave represented by these equations and the velocity of the waves is given by;

$$\frac{1}{v^2} = \mu_0\epsilon_0, \quad \text{Or}$$

$$v = \frac{1}{\sqrt{\mu_0\epsilon_0}} = C \qquad (93)$$

Substituting the values of μ_0 **and** ϵ_0, the velocity of electromagnetic waves in free space comes out to be 2.99×10^8 m/s, which is equal to the already calculated speed of light in vacuum.

As we have already shown that the variable electric and magnetic fields are coupled together. So, when we write equation of electromagnetic wave, either of the one is sufficient to consider for further treatment of the wave equation. One most important conclusion can be made here; the velocity of electromagnetic wave in free space comes out to be equation to the velocity of light in vacuum. So, in first impression, it may be said that light is an electromagnetic wave. We shall however find other physical parameters of light with these equations to finally conclude the electromagnetic nature of light. The following coupled equations having E and H, as wave parameters are called equations of electromagnetic waves in the free space.

$$\left.\begin{aligned} \nabla^2\mathbf{E} - \mu_0\epsilon_0 \frac{\partial^2\mathbf{E}}{dt^2} = 0 \\ \nabla^2\mathbf{H} - \mu_0\epsilon_0 \frac{\partial^2\mathbf{H}}{dt^2} = 0 \end{aligned}\right] \qquad (94)$$

And the velocity of electromagnetic waves in free space is calculated by;

$$C = \frac{1}{\sqrt{\mu_0\epsilon_0}}$$

For simplicity we can treat either of the two for finding various physical parameters of the electromagnetic waves.

2.26 Electromagnetic Waves in a Non-Conducting Isotropic Dielectric

Let us study the propagation of electromagnetic waves in a medium which is linear, non-conducting dielectric, isotropic and homogeneous. Let ϵ and μ are the permittivity and permeability respectively of the medium. Let the medium is source free, the electric field enters from outside the medium then the charge density ρ will be zero. Being non conducting, the conductivity $\sigma = 0$ and hence the conduction current density $\vec{J} = \sigma\vec{E}$, is also zero.

Writing, the Maxwell's equations in differential form;

$$\begin{aligned}
&\nabla.\mathbf{D} = \rho \quad \text{or} \quad \text{div}\,\mathbf{D} = \rho && \text{(i)}\\
&\nabla.\mathbf{B} = 0 \quad \text{or} \quad \text{div}\,\mathbf{B} = 0 && \text{(ii)}\\
&\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{or} \quad \text{curl}\,\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} && \text{(iii)}\\
&\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad \text{or} \quad \text{curl}\,\mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} && \text{(iv)}
\end{aligned} \qquad (95)$$

Now putting, $\mathbf{D} = \epsilon\mathbf{E}, \mathbf{B} = \mu\mathbf{H}, \mathbf{J} = \sigma\mathbf{E} = 0$ and $\rho = 0$, in (95), we have;

$$\begin{aligned}
&\nabla.\mathbf{E} = 0 && \text{(i)}\\
&\nabla.\mathbf{H} = 0 && \text{(ii)}\\
&\nabla \times \mathbf{E} = -\mu\frac{\partial \mathbf{H}}{\partial t} && \text{(iii)}\\
&\nabla \times \mathbf{H} = \epsilon\frac{\partial \mathbf{E}}{\partial t} && \text{(iv)}
\end{aligned} \qquad (96)$$

Taking curl of equation {(96)(iii)}, we get;

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu\,\frac{\partial(\nabla \times \mathbf{H})}{\partial t}$$

Simplifying the vector triple product and putting the value of $\nabla \times \mathbf{H}$ from {(96)(iv)}, we have;

$$\nabla(\nabla.\mathbf{E}) - \nabla^2\mathbf{E} = -\mu\,\frac{\partial}{\partial t}[\,\epsilon\frac{\partial \mathbf{E}}{\partial t}]$$

$$\nabla^2\mathbf{E} - \mu\epsilon\,\frac{\partial^2\mathbf{E}}{dt^2} = 0 \qquad (97)$$

Similarly, we can have following equation for magnetic field intensity, i.e.

$$\nabla^2\mathbf{H} - \mu\epsilon\,\frac{\partial^2\mathbf{H}}{dt^2} = 0 \qquad (98)$$

So, in coupled form, the equations (97) and (98) are the equations of electromagnetic waves in a linear, non-conducting dielectric, isotropic and homogeneous medium. Here, **E** and **H**, are the wave variables. Comparing (97) and (98), with the general wave equation, the speed of electromagnetic wave in isotropic, homogeneous, dielectric medium will be;

$$v = \frac{1}{\sqrt{\mu\,\epsilon}}$$

$$= \frac{1}{(K_m\mu_0K_e\epsilon_0)} \qquad (99)$$

Where K_m and K_e are the relative permeability and permittivity of the medium respectively.

As we have $C = \frac{1}{\sqrt{\mu_0\epsilon_0}}$ is the speed of electromagnetic waves in free space. Hence, the velocity of electromagnetic waves will be;

$$\therefore v = \frac{C}{\sqrt{K_mK_e}} \qquad (100)$$

As $K_m > 1$ and $K_e > 1$: therefore we can say the speed of electromagnetic waves in an isotropic dielectric is less than the speed of electromagnetic waves in free space.

The ratio of velocity of electromagnetic waves in vacuum to the medium gives another physical parameter which is widely being used for light as a relative parameter between two media called refractive index of any medium i.e.

Refractive index $= \frac{C}{v} = \sqrt{K_mK_e} = n$(say)

For a non-magnetic material $K_m = 1$; therefore, $n = \sqrt{K_e}$, i.e. $n^2 = K_e$

2.27 Transverse Nature of Electromagnetic Waves

To show the transverse nature of electromagnetic waves in a linear, non-conducting dielectric, isotropic and homogeneous medium, let us have the wave equations as derived above;

$$\nabla^2\mathbf{E} - \frac{1}{v^2}.\frac{\partial^2\mathbf{E}}{\partial t^2} = 0 \qquad (101)$$

$$\nabla^2\mathbf{H} - \frac{1}{v^2}.\frac{\partial^2\mathbf{H}}{\partial t^2} = 0 \qquad (102)$$

Mathematically these are the second order differential equations. So, the plane-wave solutions of equations (101) and (102), may be written as;

$$\mathbf{E}(r,t) = E_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} \qquad (103)$$

$$\mathbf{H}(r,t) = H_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} \qquad (104)$$

Where E_0 and H_0 are complex amplitudes which are constant in space and time, while k is wave propagation vector given by

$$\mathbf{K} = k\hat{n} = \frac{2\pi}{\lambda}\hat{n} = \frac{\omega}{v}\hat{n} \qquad (105)$$

Here $\hat{n}$ is a unit vector in the direction of wave propagation vector.

Now, as $\mathbf{E}(r,t)$ and $\mathbf{H}(r,t)$ are the solutions of the Maxwell waves equations so these should satisfy all the four Maxwell equations, i.e.

$$\nabla.\mathbf{E}(r,t) = 0 \qquad (106)$$

$$\nabla.\mathbf{H}(r,t) = 0 \qquad (107)$$

$$\nabla \times \mathbf{E}(r,t) = -\mu\frac{\partial \mathbf{H}(r,t)}{\partial t} \qquad (108)$$

$$\nabla \times \mathbf{H}(r,t) = \epsilon\frac{\partial \mathbf{E}(r,t)}{\partial t} \qquad (109)$$

Substituting $\mathbf{E}(r,t)$ and $\mathbf{H}(r,t)$ from equations (65) and (66) respectively we have;

$$\nabla.E_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} = 0 \qquad (110)$$

$$\nabla.H_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} = 0 \qquad (111)$$

Let us simplify equation (110), i.e., taking the dot product of Del operator and electric field vector we have;

$$\nabla.\mathbf{E} = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{dz}\right).E_0 e^{i\mathbf{k}.\mathbf{r}-i\omega t}$$

$$= \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{dz}\right)\left[\left(\hat{\imath}E_{0x} + \hat{\jmath}E_{0y} + \hat{k}E_{0z}\right)e^{i(k_x x+k_y y+k_z z)-i\omega t}\right]$$

$$\left[\text{since } \mathbf{k}.\mathbf{r} = \left(\hat{\imath}k_x + \hat{\jmath}k_y + \hat{k}k_z\right).(\hat{\imath}x + \hat{\jmath}y + \hat{k}z)\right]$$

$$= \left[k_x x + k_y y + k_z z\right]$$

$$\therefore \nabla.\mathbf{E} = (E_{0x}\hat{\imath}k_x + E_{0y}\hat{\imath}k_y + E_{0z}\hat{\imath}k_z)\, e^{i\mathbf{K}.\mathbf{r}-i\omega t}$$

$$= \hat{\imath}(E_{0x}k_x + E_{0y}k_y + E_{0z}k_z)\, e^{i\mathbf{K}.\mathbf{r}-i\omega t}$$

$$= \hat{\imath}\left(\hat{\imath}k_x + \hat{\jmath}k_y + \hat{k}k_z\right).(\hat{\imath}E_{0x} + \hat{\jmath}E_{0y} + \hat{k}E_{0z})\, e^{i\mathbf{K}.\mathbf{r}-i\omega t}$$

$$= \hat{\imath}\mathbf{K}.E_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} \quad = \hat{\imath}\mathbf{K}.\mathbf{E} = 0$$

Thus $\nabla.\mathbf{E} = 0$ implies that

$$\mathbf{K}.\mathbf{E} = 0 \qquad (112)$$

Now, $\mathbf{K}.\mathbf{E} = 0$, implies the wave propagation vector **K** and the electric field vector **E** of the electromagnetic wave are perpendicular to each other.

On doing similar operation on equation (111), we have;

$$\nabla.\mathbf{H} = \hat{\imath}\,\mathbf{K}.\mathbf{H}$$

Or

$$\mathbf{K}.\mathbf{H} = 0 \qquad (113)$$

Again, we can infer the propagation vector **K** and the oscillating magnetic field vector **H** are also perpendicular to each other. The above mathematical operations, shows that both the wave variables **E** and **H** oscillate perpendicular to the propagation of wave. This proves the transverse character of electromagnetic waves.

2.28 Mutually Perpendicularity of Electric and Magnetic Field Vectors

On satisfying the equations (101) and (102) with the solutions given by (103) and (104), we have;

$$\nabla \times \mathbf{E}(r,t) = -\mu \frac{\partial \mathbf{H}(r,t)}{\partial t} \quad (114)$$

$$\nabla \times \mathbf{H}(r,t) = \epsilon \frac{\partial \mathbf{E}(r,t)}{\partial t} \quad (115)$$

$$\nabla \times E_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} = -\mu \frac{\partial H_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t}}{\partial t} \quad (116)$$

$$\nabla \times H_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t} = \epsilon \frac{\partial E_0 e^{i\mathbf{K}.\mathbf{r}-i\omega t}}{\partial t} \quad (117)$$

On simplifying the mathematical operations of the equations (116) and (117), we find;

$$i\mathbf{K} \times \mathbf{E} = i\mu\omega\mathbf{H} \quad \text{i.e. } \mathbf{K} \times \mathbf{E} = \mu\omega\mathbf{H} \quad (118)$$

$$\mathbf{K} \times \mathbf{H} = -\epsilon\omega\mathbf{E} \quad (119)$$

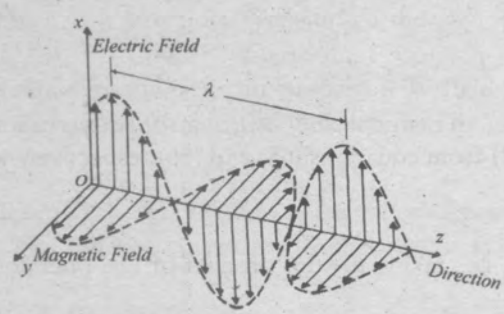


Fig. 2.10 Transverse Character of Electromagnetic wave

The equations (118) and (119) shows that the wave propagation vector **k**, electric field vector **E** and Magnetic field vector **H** are mutually perpendicular to each other. Simply we can say the electric and magnetic fields are oscillating perpendicular to each other.

2.29 Some Other Physical Parameters of Electromagnetic Wave

By doing other mathematical manipulations on the equations derived above we are able to define a lot of physical quantities related to the electromagnetic waves.

2.29.1 Wave Impedance

On rearranging the equation (118), we have;

$$\mathbf{H} = \frac{1}{\mu\omega}(\mathbf{K} \times \mathbf{E}) = \frac{k}{\mu\omega}(\hat{n} \times \mathbf{E})$$

Or

$$\left|\frac{E}{H}\right| = \frac{\mu\omega}{K} = \mu v, \qquad \left(\text{since } \mathbf{K} = \frac{\omega}{v}\right)$$

Or

$$\left|\frac{E}{H}\right| = \frac{\mu\omega}{K} = \frac{\mu}{\sqrt{\mu\epsilon}} \qquad \left(\text{since } v = \frac{1}{\sqrt{\mu\epsilon}}\right)$$

Or

$$\left|\frac{E}{H}\right| = \sqrt{\frac{\mu}{\epsilon}} \quad (120)$$

Now let us see what this quantity is showing. Putting the units of each term involved in it, we have;

$$\frac{\text{volt/m}}{\text{amp-turn/m}} = \frac{\text{volt}}{\text{amp}} = \text{ohm}$$

This implies that the magnitude of $\left|\frac{E}{H}\right| = \sqrt{\frac{\mu}{\epsilon}}$ is giving some physical quantity which has the unit as ohm. This unit represents impedance. Here, this is called wave impedance offered by medium when electromagnetic waves propagate through it and is given by the modulus of the ratio of electric field vector and magnetic field vector. This can also be calculated by the square root of the ratio of permeability and permittivity of the medium. So, the wave impedance is given as;

$$Z = \left|\frac{E}{H}\right| = \sqrt{\frac{\mu}{\epsilon}}$$

As we have the values of the permeability and permittivity of the free space as μ_0 and ϵ_0 , the impedance offered by the vacuum when electromagnetic waves propagate through it is given by;

$$Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$$

Substituting the values of μ_0 and ϵ_0, the value of Z_0 comes out to be as;

$$Z_0 = 376.66 \text{ ohm}$$

2.29.2 Phase between Electric Field and Magnetic Field Vector

A simple way to find whether the two vector quantities are in phase or out of phase is that if the modulus of the ratio of the two quantities is real then those two vectors are said to be in phase but if the modulus of the ratio, is a complex quantity then the two vector quantities are said to be out of phase. As we already have the modulus of the ratio of E and H in free space as well as in isotropic dielectric, i.e.

$$Z_0 = \left|\frac{E_0}{H_0}\right| = \sqrt{\frac{\mu_0}{\epsilon_0}} = 376.6 \text{ ohm, and } Z = \left|\frac{E}{H}\right| = \sqrt{\frac{\mu}{\epsilon}} = \left(\frac{K_m \mu_0}{K_e\, \epsilon_0}\right)$$

Both of these are real. So, we can say when electromagnetic waves propagate in free space and isotropic dielectric, the oscillations of electric field and magnetic field are in phase with each other.

2.29.3 Poynting Vector for a Plane Electromagnetic Wave in an Isotropic Dielectric

Let us find out the flow of electromagnetic energy when electromagnetic waves propagate through an isotropic dielectric. The well-defined way of doing this is to find out the poynting vector of electromagnetic waves in corresponding medium.

The poynting vector is calculated as;

$$\begin{aligned}
\mathbf{S} &= \mathbf{E} \times \mathbf{H} \\
&= \mathbf{E} \times \left(\frac{K}{\mu\omega}(\hat{n} \times \mathbf{E})\right) \\
&= \mathbf{E} \times (\hat{n} \times \mathbf{E}).\frac{1}{\mu v} \\
&= \mathbf{E} \times (\hat{n} \times \mathbf{E}).\sqrt{\frac{\epsilon}{\mu}} \\
&= \frac{\mathbf{E} \times (\hat{n} \times \mathbf{E})}{Z} && \left(\text{since } Z = \sqrt{\frac{\mu}{\epsilon}}\right) \\
&= \frac{(\mathbf{E}.\mathbf{E})\hat{n} - (\mathbf{E}.\hat{n})\mathbf{E}^2}{Z} \\
&= \frac{E^2}{Z}\hat{n} && (\text{since } E.\hat{n} = 0, \text{ as E and } \hat{n} \text{ are perpendicular}) \qquad (121)
\end{aligned}$$

As **E** and **H** are oscillating between maximum and minimum values of these physical quantities, so to find the flow of energy of electromagnetic waves, we have to take time average of poynting vector over a complete cycle of electric of magnetic field vector i.e.;

$$\langle \mathbf{S} \rangle = \langle \mathbf{E} \times \mathbf{H} \rangle = \langle \frac{E^2}{Z} \hat{n} \rangle$$
$$= \frac{1}{Z} \langle \left(E_0 e^{i\mathbf{K}.\mathbf{r} - i\omega t}\right)^2_{real} \rangle \hat{n}$$

Since for finding actual physical fields, we often take real parts of complex exponentials. So, we have;

$$\langle \mathbf{S} \rangle = \frac{1}{Z} E_0{}^2 \langle \cos^2(\omega t - \mathbf{K}.\mathbf{r}) \rangle \hat{n}$$
$$= \frac{1}{Z} E_0{}^2 . \frac{1}{2} \hat{n} = \frac{E_0{}^2}{2Z} \hat{n}$$
$$= \frac{E_{rms}^2}{Z} \hat{n} \qquad \left(\text{since } E_{rms} = \frac{E_0}{\sqrt{2}}\right) \qquad (122)$$

The equation (122) gives the average energy flow of electromagnetic waves, per unit area per time i.e., energy flux. It also shows that the flow of energy is along the direction of propagation of electromagnetic wave.

Similarly for propagation in free space we have the time average poynting vector as;

$$\langle S \rangle_{\text{free space}} = \frac{E_{rms}^2}{Z_0} \hat{n} \qquad (123)$$

2.29.4 Power Flow and Energy Density

Let us find the ratio of electrostatic and magnetostatic energy densities in an electromagnetic wave field i.e.

$$\frac{u_e}{u_m} = \frac{\frac{1}{2} \epsilon E^2}{\frac{1}{2} \mu H^2} = \frac{\epsilon}{\mu} \frac{E^2}{H^2} = \frac{\epsilon}{\mu} Z^2 = \frac{\epsilon}{\mu} . \frac{\mu}{\epsilon} = 1 \qquad (124)$$

This implies that for the case of electromagnetic waves in an isotropic dielectric the electrostatic energy density (u_e) is equal to the magnetostatic energy density (u_m).

Therefore, total electromagnetic energy density

$$u = u_e + u_m = 2u_e \; (\text{since } u_e = u_m)$$
$$= 2.\frac{1}{2} \epsilon E^2 = \epsilon E^2$$

And the time average of energy density

$$\langle u \rangle = \epsilon \langle E^2 \rangle = \epsilon \langle \left(E_0 e^{i\mathbf{K}.\mathbf{r} - i\omega t}\right)^2_{real} \rangle$$
$$= \epsilon E^2 \langle \cos^2(\omega t - \mathbf{K}.\mathbf{r}) \rangle$$
$$= \frac{1}{2} \epsilon E_0^2$$

i.e., the total electromagnetic energy density $\langle u \rangle = \frac{1}{2} \epsilon E_0^2 = \epsilon E_{rms}^2$ (125)

2.29.5 Relation between Energy Flux and Energy Density of Electromagnetic Waves

Dividing equation (123) and (125), we have;

$$\frac{\langle S \rangle}{\langle u \rangle} = \frac{E_{rms}^2 \hat{n}/Z}{\epsilon E_{rms}^2} = \frac{1}{\epsilon Z} \hat{n} = \frac{1}{\sqrt{\frac{\mu}{\epsilon}} \epsilon} \hat{n} \qquad (\text{since } Z = \sqrt{\frac{\mu}{\epsilon}})$$
$$= \frac{1}{\sqrt{\mu \epsilon}} \hat{n} = v\hat{n} \qquad \text{As } v = \frac{1}{\sqrt{\mu \epsilon}}$$
$$\frac{\langle S \rangle}{\langle u \rangle} = v\hat{n}$$
$$\langle S \rangle = \langle u \rangle v\hat{n} \qquad (126)$$

Energy Flux = Energy Density × velocity of em waves in the medium
This equation implies that the energy density associated with an electromagnetic wave in the dielectrics, flows with the same speed of the wave in the dielectrics and it is in the direction of propagation of the wave.

Following points can be summarised for the electromagnetic waves in isotropic dielectric:

- The electromagnetic waves travel with a speed less than the speed of light in case of an isotropic dielectric.
- The electromagnetic waves are transverse in nature as the field vectors E and H are mutually perpendicular and also perpendicular to the direction of propagation the electromagnetic wave.
- The phase of field vectors E and H is same.
- The direction of flow of electromagnetic energy and the direction of wave propagation is same and the energy flowing per second is represented as $\langle S\rangle \frac{E_{rms}^2\,\hat{n}}{Z} = \langle u\rangle v\hat{n}$
- The electrostatic energy density and the magnetostatic energy density are equal and the total energy density is given as $\langle u\rangle = \epsilon\, E_{rms}^2$
- The energy density associated with an electromagnetic wave in dielectrics flows with the same speed as of the wave in dielectric and in the direction of propagation of the wave.

2.30 Plane Electromagnetic Waves in a Conducting Medium

Maxwell equations in differential form are;

$$\text{div}\,\mathbf{D} = \nabla.\mathbf{D} = \rho$$
$$\text{div}\,\mathbf{B} = \nabla.\mathbf{B} = 0$$
$$\text{curl}\,\mathbf{E} = \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}$$
$$\text{curl}\,\mathbf{H} = \nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial\mathbf{D}}{\partial t}$$

Let us assume that medium is linear and isotropic and is characterised by permittivity ϵ permeability μ and conductivity σ, but not any charge or any current other than that determined by Ohm's law. Then the parameters of the medium are $\mathbf{D} = \epsilon\mathbf{E}, \mathbf{B} = \mu\mathbf{H}, \mathbf{J} = \sigma\mathbf{E}$ and $\rho = 0$

So that Maxwell's equations for a linear, isotropic, homogeneous and conducting medium will be;

$$\left.\begin{array}{ll} \nabla.\mathbf{E} = 0 & \text{(i)} \\ \nabla.\mathbf{H} = 0 & \text{(ii)} \\ \nabla\times\mathbf{E} = -\mu\frac{\partial\mathbf{H}}{\partial t} & \text{(iii)} \\ \nabla\times\mathbf{H} = \sigma\mathbf{E} + \epsilon\frac{\partial\mathbf{E}}{\partial t} & \text{(iv)} \end{array}\right\} \quad (127)$$

Taking curl of equation {(127),(iii)}, we get;

$$\nabla\times(\nabla\times\mathbf{E}) = -\mu\frac{\partial}{\partial t}(\nabla\times\mathbf{H}) \quad (128)$$

Substituting $(\nabla\times\mathbf{H})$ from equation{(127) (iv)} into (128), we get;

$$\nabla\times(\nabla\times\mathbf{E}) = -\mu\frac{\partial}{\partial t}\left(\sigma\mathbf{E} + \epsilon\frac{\partial\mathbf{E}}{\partial t}\right)$$

$$\nabla\times(\nabla\times\mathbf{E}) = -\sigma\mu\frac{\partial\mathbf{E}}{\partial t} - \epsilon\mu\frac{\partial^2\mathbf{E}}{\partial t^2} \quad (129)$$

As we know, $\nabla\times(\nabla\times\mathbf{E}) = \nabla(\nabla.\mathbf{E}) - \nabla^2\mathbf{E}$ (130)

Using identity (130) and putting $\nabla.\mathbf{E} = 0$ for the medium, the equation (129) takes the form;

$$\nabla^2\mathbf{E} = -\sigma\mu\frac{\partial\mathbf{E}}{\partial t} - \epsilon\mu\frac{\partial^2\mathbf{E}}{\partial t^2} \quad (131)$$

Similarly doing similar operation for magnetic field, we get;

$$\nabla^2 \mathbf{H} = -\sigma\mu\frac{\partial \mathbf{H}}{\partial t} - \epsilon\mu\frac{\partial^2 \mathbf{H}}{\partial t^2} \tag{132}$$

The equations (131) and (132) represent the equations of electromagnetic wave in a linear, homogeneous, isotropic, conducting medium of conductivity σ.

The general solutions of these second order differential equation (131) and (132)), is written as;

$$\mathbf{E} = E_0 e^{i\mathbf{K}.\mathbf{r} - i\omega t} \tag{133}$$

$$\mathbf{H} = H_0 e^{i\mathbf{K}.\mathbf{r} - i\omega t} \tag{134}$$

Where the wave vector, **K**, may be complex, while E_0 and H_0 are complex amplitudes which are constant in space and time. Satisfying the solutions from (133) and (134) with the corresponding differential equation (131) and (132) respectively, we get;

$$(-k^2 + i\sigma\mu\omega + \mu\epsilon\omega^2)\mathbf{E} = 0$$

$$(-k^2 + i\sigma\mu\omega + \mu\epsilon\omega^2)\mathbf{H} = 0$$

As the fields **E** or **H** are arbitrary, therefore this equation holds only if

$$(-k^2 + i\sigma\mu\omega + \mu\epsilon\omega^2) = 0$$

$$k^2 = \mu\epsilon\omega^2\left(1 + \frac{i\sigma}{\omega\epsilon}\right) = \mu\epsilon\omega^2 + i\,\sigma\mu\omega \tag{135}$$

From above equation, it seems that the wave propagation vector **K** is a complex quantity in the case, when the electromagnetic waves propagate in a linear, isotropic, homogeneous and conducting medium.

2.30.1 Other Physical Parameters of the Wave

2.30.1.1 Propagation Vector

Let us further simplify the equation (135) to find out some more physical parameters of electromagnetic waves in a conducting medium.

Let $\mathbf{K} = \alpha + i\beta$

So $\mathbf{K}^2 = \alpha^2 - \beta^2 + 2i\alpha\beta$ (136)

Comparing equation (97) and (98), we have

$$\alpha^2 - \beta^2 = \mu\epsilon\omega^2 \tag{137}$$

And

$$2\alpha\beta = \mu\omega\sigma \tag{138}$$

Or

$$\beta = \frac{\mu\omega\sigma}{2\alpha},$$

So, we have;

$$\alpha^2 - \left(\frac{\mu\omega\sigma}{2\alpha}\right)^2 = \mu\epsilon\omega^2$$

$$\alpha^4 - \mu\epsilon\omega^2\alpha^2 - \frac{\mu^2\omega^2\sigma^2}{4} = 0$$

$$(\alpha^2)^2 - \mu\epsilon\omega^2\alpha^2 - \frac{\mu^2\omega^2\sigma^2}{4} = 0$$

Comparing the above equation with the standard quadratic equation, we have the coefficients, a, b and c as;

$$a = 1,\ b = -\mu\epsilon\omega^2, \quad c = -\frac{\mu^2\omega^2\sigma^2}{4}$$

$$\alpha^2 = \mu\epsilon\omega^2 \pm \frac{\sqrt{\mu^2\epsilon^2\omega^4 + 4\times 1\times\frac{\mu^2\omega^2\sigma^2}{4}}}{2}$$

$$= \frac{\mu\epsilon\omega^2 \pm \sqrt{\mu^2\epsilon^2\omega^4 + \mu^2\omega^2\sigma^2}}{2}$$

$$= \frac{\mu\epsilon\omega^2 \pm \sqrt{\mu^2\epsilon^2\omega^4\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]}}{2}$$

$$\alpha = \sqrt{\mu\epsilon}\,.\,\omega\left[\frac{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}+1}{2}\right]^{1/2} \tag{139}$$

Similarly, the value of β comes out to be;

$$\beta = \sqrt{\mu\epsilon}\,.\,\omega\left[\frac{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}-1}{2}\right]^{1/2} \tag{140}$$

So, the wave propagation vector K of an electromagnetic wave travelling in a conducting medium having conductivity σ , μ permeability, ϵ permittivity and ω as the frequency of em waves, can be calculated as;

$$\mathbf{K} = \sqrt{\mu\epsilon}\,.\,\omega\left[\frac{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}+1}{2}\right]^{1/2} + i\sqrt{\mu\epsilon}\,.\,\omega\left[\frac{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}-1}{2}\right]^{1/2} \tag{141}$$

Which is a complex quantity.

Case 1. Propagation vector for Very Good Conductor

For a good conductor $\frac{\sigma}{\omega\epsilon} \gg 1$, so 1 can be neglected in comparison of $\frac{\sigma}{\omega\epsilon}$ so that α and β are approximately equal i.e.

$$\alpha \approx \beta = \sqrt{\mu\epsilon}.\,\omega\sqrt{\frac{\frac{\sigma}{\omega\epsilon}}{2}} = \sqrt{\frac{\mu\sigma\omega}{2}}$$

So, the propagation constant comes out to be

$$\mathbf{K} \approx \alpha + i\beta = \sqrt{\frac{\mu\sigma\omega}{2}} + i\sqrt{\frac{\mu\sigma\omega}{2}}$$

Case 2. Propagation vector for Very Poor Conductor

For very poor conductor, $\frac{\sigma}{\omega\epsilon} \ll 1$, and be neglected then

$\alpha = \sqrt{\mu\epsilon}.\,\omega$ and β may neglected or can be taken as zero.

$$\therefore\ \mathbf{K} = \alpha + i\beta = \sqrt{\mu\epsilon}.\,\omega \tag{142}$$

Or $\mathbf{K} = \frac{\omega}{v}$, since $v = \frac{1}{\sqrt{\mu\epsilon}}$

This comes out to be as in the case of non-conductor and also shows negligible or zero attenuation of electromagnetic waves in a very poor conductor, i.e., may be treated as a homogeneous isotropic non-conducting medium.

2.30.1.2 Attenuation of Electromagnetic Waves in a Conducting Medium

Now putting the value of $\mathbf{K} = \alpha + i\beta$, in the equations (95) and (96) we have:

$$E = E_0 e^{i(\alpha+i\beta)\,\mathbf{n.r} - i\omega t} = E_0 e^{-\beta \mathbf{n.r}} e^{i\,(\alpha \mathbf{n.r} - \omega t)} \qquad (143)$$

$$H = H_0 e^{i(\alpha+i\beta)\,\mathbf{n.r} - i\omega t} = H_0 e^{-\beta \mathbf{n.r}} e^{i\,(\alpha \mathbf{n.r} - \omega t)} \qquad (144)$$

From equations (133) and (134), it can be seen that field amplitudes of the electric and magnetic oscillations i.e., $E_0 e^{-\beta \mathbf{n.r}}$ and $H_0 e^{-\beta \mathbf{n.r}}$ respectively, are attenuated with distance due to the term $e^{-\beta \mathbf{n.r}}$. We can say the amplitude of electric and magnetic fields vectors decay exponentially. The quantity β is a characteristic of the medium and is known as decay constant or absorption coefficient. So, the attenuation coefficient of any electromagnetic wave in a conducting medium is given by;

$$\beta = \sqrt{\mu\epsilon}\,.\,\omega \left[\frac{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}-1}{2}\right]^{1/2} \qquad (145)$$

2.30.1.3 Penetration of Electromagnetic Wave into a Medium - skin Depth or Penetration Depth

Now, we have seen that as the electromagnetic waves travel through a conducting medium, the amplitude of the field vectors, attenuate with distance. The distance/depth from the surface of the medium, at which the amplitude becomes $\frac{1}{e} = 0.369$, of its initial amplitude at the surface, i.e., $E = \frac{1}{e} E_0$, is called skin depth or penetration depth.

For this, the distance r should be $\frac{1}{\beta}$ and is denoted by δ.

So, the penetration of skin depth is given by;

$$\delta = r = \frac{1}{\beta} = \frac{1}{\sqrt{\mu\epsilon}\,.\omega} \left[\frac{2}{\sqrt{\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}}-1}\right]^{1/2} \qquad (146)$$

For good conductor, the penetration depth is given by

$$\delta = \frac{1}{\beta} = \sqrt{\frac{2}{\mu\sigma\omega}} \qquad (147)$$

This expression may be useful to find the thickness of a conducting enclosure for protection from electromagnetic radiation. This also shows that very high frequency electromagnetic waves travel through the surface of the conductors.

Let us take a few examples:

(i) For copper at 60 cycles δ is 0.86 cm, but at 1 megacycle, it has dropped to 0.0067. That is why in high frequency circuits current flows only on the surface of the conductors. The major importance of the skin depth is that it measures the depth to which an electromagnetic wave can penetrate a conducting medium. Therefore, the conducting sheets which are used as electromagnetic shields must be thicker than the skin depth.

(ii)) For silver $\sigma \approx 10^7 \text{mho/m}$ at a typical microwave frequency$\approx 10^8 \frac{c}{s}$, the skin depth $\approx 10^{-4}$ cm. Thus at microwave frequencies the skin depth in silver is very small and consequently performance of a pure silver component and a silver-plated brass component would be expected to be indistinguishable.

(iii) For sea water $\sigma \approx 4.3\ \text{mho/m}$ at a frequency of 60 kc/s ; so that $\delta \approx 1\text{meter}$. That is why radiocommunication with submerged submarine becomes increasing difficult at several skin depths.

2.31 Transverse Character of Electromagnetic Waves in Conducting Medium

Now satisfying the general solutions of differential equations of electromagnetic waves, we get;

$$i\mathbf{K}.\mathrm{E} = 0 \quad \text{or} \quad \mathbf{K}.\mathrm{E} = 0 \tag{148}$$

$$i\mathbf{K}.\mathrm{H} = 0 \quad \text{or} \quad \mathbf{K}.\mathrm{H} = 0 \tag{149}$$

These equations imply that field vectors **E** and **H** are both perpendicular to the direction of propagation vector **K**. This implies that electromagnetic waves in a conducting medium are transverse in nature.

On satisfying third and fourth equations we get

$$\nabla \times \mathbf{E}(\mathrm{r,t}) = -\mu \frac{\partial \mathbf{H}(\mathrm{r,t})}{\partial t} \tag{150}$$

$$\nabla \times \mathbf{H}(\mathrm{r,t}) = \epsilon \frac{\partial \mathbf{E}(\mathrm{r,t})}{\partial t} \tag{151}$$

$$\nabla \times \mathrm{E_0} e^{i\mathbf{K}.\mathbf{r}-i\omega t} = -\mu \frac{\partial \mathrm{H_0} e^{i\mathbf{K}.\mathbf{r}-i\omega t}}{\partial t} \tag{152}$$

$$\nabla \times \mathrm{H_0} e^{i\mathbf{K}.\mathbf{r}-i\omega t} = \epsilon \frac{\partial \mathrm{E_0} e^{i\mathbf{K}.\mathbf{r}-i\omega t}}{\partial t} \tag{153}$$

On simplification (152) and (153), we get;

$$i\mathbf{K} \times \mathbf{E} = i\mu\omega\mathbf{H} \text{ i.e. } \mathbf{K} \times \mathbf{E} = \mu\omega\,\mathrm{H} \tag{154}$$

and

$$i\mathbf{K} \times \mathbf{H} = (\sigma - i\epsilon\omega)\mathbf{E} \text{ ie. } \mathbf{K} \times \mathbf{H} = (\epsilon\omega + i\sigma)\mathbf{E} \tag{155}$$

The equations (154) and (155), imply that the electromagnetic field vectors **E** and **H** are mutually perpendicular and also they are perpendicular to the direction of propagation vector **K**, in a conducting medium also.

2.32 Relative Phase of E and H

Using equations (154) and (155), we have;

$$\mathbf{H} = \frac{1}{\mu\omega}(\mathbf{K} \times \mathbf{E}) = \frac{1}{\mu\omega}\mathrm{k}(\mathrm{n} \times \mathbf{E})$$

$$= \frac{(\alpha+i\beta)}{\mu\omega}(\mathrm{n} \times \mathbf{E}) \tag{156}$$

This implies that $\left|\frac{\mathrm{H}}{\mathrm{E}}\right| = \frac{\mathrm{H_0}}{\mathrm{E_0}} = \frac{\alpha+i\beta}{\mu\omega} = \text{complex quantity}$ (157)

This implies that the field vectors H and E are out of phase in a conductor. The magnitude and phase of complex wave vector is written as $= |\mathrm{k}|e^{i\phi}$. and can be calculated as;

$$|\mathbf{K}| = |\alpha + i\beta| = \sqrt{(\alpha^2 + \beta^2)} = \sqrt{\mu\epsilon}\omega\left[1 + \left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/4} \tag{158}$$

$$\text{and } \phi = \tan^{-1}\left(\frac{\beta}{\alpha}\right) = \frac{1}{2}\tan^{-1}\left(\frac{\sigma}{\omega\epsilon}\right) \tag{159}$$

So equation (156) may be expressed as

$$\mathrm{H} = \frac{1}{\mu\omega}\sqrt{\mu\epsilon}\omega\left[1 + \left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} e^{i\phi}\,(\mathrm{n} \times \mathrm{E})$$

$$= \frac{\epsilon}{\mu}\left[1 + \left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} e^{i(\phi)}\,(\mathrm{n} \times \mathrm{E}) \tag{160}$$

This equation shows that H lags behind E by the phase angle ϕ given by equation (159) and the relative magnitude of magnetic and electric field is;

$$\left|\frac{H}{E}\right| = \frac{H_0}{E_0} = \sqrt{\frac{\epsilon}{\mu}}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} = \frac{1}{Z}, \text{ Where Z is the wave impedence in the conducting medium.} \quad (161)$$

2.33 Poynting Vector or Energy Flow of Electromagnetic Waves in Conducting Medium

The Poynting vector is given by:

$\mathbf{S = (E \times H)}$

And the time average of poynting vector may expressed as

$$S_{av} = \frac{1}{2}\mathrm{Re}\left[E \times \left\{\frac{\epsilon}{\mu}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/2} e^{i(-\phi)}\,(n \times E)\right\}\right]$$

$$= \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} \mathrm{Re}\{E \times (n \times E^*)e^{i(-\phi)}\}$$

$$= \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} \mathrm{Re}\left[\{(E.E^*)n) - (E.n)E^*\}e^{i(-\phi)}\right]$$

$$= \frac{1}{2}\sqrt{\frac{\epsilon}{\mu}}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{\frac{1}{4}} E_0^2\, e^{-2\beta n.r} n \cos\emptyset \quad (162)$$

$\left[\text{Since } (E.E^*) = E_0^2\, e^{-2\beta n.r} \text{ and } \mathrm{Re}\,(e^{i(-\phi)} = \cos\emptyset\right]$

For good conductor $\sigma/\epsilon\omega \gg 1$ so that $\emptyset = \pi/4$ and also $E_{rms} = \frac{E_0}{\sqrt{2}}$

Hence $S_{av} = \left\{\left(\frac{\sigma}{2\mu\omega}\right)\right\} E_{rms}^2\; e^{-2\beta n.r}$

Energy density: The total energy density of electromagnetic field is given by

$$u = u_e + u_m$$

where electrostatic energy u_e

$$= \frac{1}{2}\,\mathrm{Re}\frac{1}{2}(E.D^*)$$

$$= \frac{1}{2}\epsilon E_0^2 e^{-2\beta n.r} \quad (163)$$

$$= \frac{1}{2}\epsilon E^2_{rms\; e^{-2\beta n.r}} \quad (164)$$

and magnetic energy density $u_m = \frac{1}{2}\;\mathrm{Re}\frac{1}{2}(H.B^*)$

$$= \frac{1}{4}\; u\,\mathrm{Re}(H.H^*)$$

$$= \frac{1}{4}\;\mu\, H_0^2\; e^{-2\beta n.r}$$

$$= \frac{1}{4}\;\mu\frac{\epsilon}{\mu}\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/2} E_0^2\; e^{-2\beta n.r} \quad \text{using (161)}$$

$$= \frac{1}{4}\epsilon\left\{\left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/2} E_0^2\; e^{-2\beta n.r}\right\}$$

$$= \left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/2} u_e \quad (165)$$

So, the total energy density $u = u_e + u_m = u_e + \left[1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right]^{1/2} u_e$

$$= \left[1+\left\{1+\left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}\right]^{1/2} u_e$$

$$= \left[1 + \left\{1 + \left(\frac{\sigma}{\omega\epsilon}\right)^2\right\}\right]^{1/2} \times \frac{1}{2}\epsilon E_{rms}^2\ e^{-2\beta \mathbf{n}.\mathbf{r}} \qquad (166)$$

From equations (162) and (166) it is clear that the energy flux and energy density are damped as the electromagnetic wave propagates in a conducting medium. This energy lose is due to Joule heating of the medium.

Following points can be summarised for the electromagnetic waves in conducting medium

- The electromagnetic waves are transverse in nature in which the electric field vector E and magnetic field vector H are mutually perpendicular and perpendicular to the direction of propagation of electromagnetic wave.
- The amplitudes of electric and magnetic field vectors E and H respectively are damped exponentially as the wave propagates deeper in the conductors.
- The electric and magnetic field vectors E and H of the electromagnetic wave are not in the same phase as H lags behind E by angle ∅ given by
$$\emptyset = \frac{1}{2}\tan^{-1}\left(\frac{\sigma}{\epsilon\omega}\right)$$
- And the magnitude of H is much greater that than of E.
- The energy flow is along the direction of propagation of electromagnetic wave and is damped exponentially as the wave propagates in the conducting medium.
- The magnetic energy density is much greater than electric energy density and both are damped exponentially as the wave propagates in the conducting medium.

******Solved examples******

Based on S.H.M

Ex.1 If a particle moves in a potential energy field $U = U_0 - ax + bx^2$ where a, and b are positive constants, obtain an expression for the force acting on it as a function of position. At what point does the force vanish? Is this point of stable equilibrium? Calculate the force constant, time period and frequency of the particle.

Sol. Force acting on the particle is given as

(i) $F = -\frac{dU}{dx} = -\frac{d}{dx}(U_0 - ax + bx^2) = a - 2bx$

(ii) The force vanishes at the point where $\frac{dU}{dx} = 0$ i.e.

$$a - 2bx = 0 \text{ or } \qquad x = a/2b$$

Which gives the position of point where the force vanishes

(iii) $\frac{d^2U}{dx^2} = 2b$ such that b is positive, the point x = a/2b represents the point of minimum potential energy on the energy curve of the particle. It is point of stable equilibrium

(iv) From the expression of force F in (i)above, the linear restoring force and force constant k is equal to 2b.

Ex.2 What is the frequeney of a simple pendulum 2.0 meters long?

Sol. The time period of a simple pendulum is given by:

$$T = 2\pi\sqrt{\frac{l}{g}}$$

Frequency,

$$n = 1/T = \frac{1}{2\pi}\sqrt{\frac{g}{l}}$$

Here $l = 2m$, $g = 9.8 m/s^2$

Frequency

$$n = \frac{1}{T} = \frac{1}{2\pi}\sqrt{\frac{9.8}{2}} = \frac{1}{2\pi}\sqrt{4.9} = 0.3524/sec$$

Based on gradient, divergence, curl

Ex.1 A scalar field u is given by $u = x^3y - xz^2 + yz$. Find grad u and its value at the point (0, 2,-1).

Sol. We have $u = x^3y - xz^2 + yz$

$\therefore$ grad $u = \nabla u = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)[x^3y - xz^2 + yz]$

$= \mathbf{i}(3x^2y - z^2) + \mathbf{j}(x^2+z) + \mathbf{k}(+2x+y)$

This is grad u. It's value at (0, 2, -1)is given as

$(\text{grad } u)_{(0,2,-1)} = \mathbf{i}(-1) + \mathbf{j}(-1) + \mathbf{k}(2)$

$= -\mathbf{i} - \mathbf{j} + 2\mathbf{k}$

Ex.2 Find grad $\mathbf{r}^n$ where r is a position vector.

Sol. We have $\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ and $r = |r| = \sqrt{x^2 + y^2 + z^2}$

$\therefore r^n = (x^2 + y^2 + z^2)^{n/2}$

$\therefore r^n = \nabla r^n = \left[\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right](x^2 + y^2 + z^2)^{n/2}$

$= \mathbf{i}\frac{n}{2}(x^2 + y^2 + z^2)^{(n/2)-1}.2x + \mathbf{j}.\frac{n}{2}(x^2 + y^2 + z^2)^{(n/2)-1}\,2y + \mathbf{k}.\frac{n}{2}(x^2 + y^2 + z^2)^{(n/2)-1}.2z$

$= (\mathbf{i}x + \mathbf{j}y + \mathbf{k}z)\left[n(x^2 + y^2 + z^2)^{(n/2)-1}\right]$

$= \cancel{(\mathbf{r}.(nr^2)^{(n/2)-1}} = r.n\,r^{n-2} = r.n\,(r^{2(\frac{n}{2}-1)})$

Thus, grad $r^2 = nr^{n-2} = nr^{n-1}\hat{r}$

In particular, if $n \mp -1$ or, respectively, we have grad $r = \hat{r}$

And grad $\frac{1}{r} = -\frac{\hat{r}}{r^2}$. We can obtain these results by direct calculation also.

Ex.3 Show that grad log $\mathbf{r} = \frac{\hat{r}}{r}$

Sol. We have

grad log $r = \nabla \log r = \left(\mathbf{i}\frac{\partial V}{\partial x} + \mathbf{j}\frac{\partial V}{\partial y} + \mathbf{k}\frac{\partial V}{\partial z}\right)\log\left(\sqrt{x^2 + y^2 + z^2}\right)$

$= \mathbf{i}\,\frac{1}{\sqrt{x^2+y^2+z^2}}.\frac{1}{2}(x^2 + y^2 + z^2)^{-1/2}.2x + \mathbf{j}.\frac{1}{\sqrt{x^2+y^2+z^2}}.\frac{1}{2}(x^2 + y^2 + z^2)^{-1/2}.2y + \mathbf{k}.$

$\frac{1}{\sqrt{x^2+y^2+z^2}}.\frac{1}{2}(x^2 + y^2 + z^2)^{-1/2}.2z$

$= \frac{\mathbf{i}x + \mathbf{j}y + \mathbf{k}z}{x^2+y^2+z^2} = \frac{\mathbf{r}}{r^2} = r\frac{\hat{r}}{r} = \frac{\hat{r}}{r}$

Thus, grad log r = $\frac{\hat{r}}{r}$

Ex.4 A spherical equipotential surface is given by

$V(x, y, z) = a(x^2 + y^2 + z^2)$

Where a is a constant. Show that the force field is radial.

Sol. We know

$$F = -\text{ grad } V = [\frac{\partial V}{\partial x}\hat{\mathbf{i}} + \frac{\partial V}{\partial y}\hat{\mathbf{j}} + \frac{\partial V}{\partial z}\hat{\mathbf{k}}]$$

$$\frac{\partial V}{\partial x} = 2ax,\ \frac{\partial V}{\partial y} = 2ay \text{ and } \frac{\partial V}{\partial z} = 2az$$

$$F = -2a\,[x\,\hat{\mathbf{i}} + y\,\hat{\mathbf{j}} + z\,\hat{\mathbf{k}}] = -2ar$$

Where $\mathbf{r} = x\,\hat{\mathbf{i}} + y\,\hat{\mathbf{j}} + z\,\hat{\mathbf{k}}$ is the position vector of point (x, y, z). since $F \propto r$, it is a radial force field.

Ex.5 If **r** is the position vector ($\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$), evaluate div r and div $\hat{r}$.

Sol. We have

$$\text{div } r = \frac{\partial r_x}{\partial x} + \frac{\partial r_y}{\partial y} + \frac{\partial r_z}{\partial z} = \frac{\partial r}{\partial x} + \frac{\partial r}{\partial y} + \frac{\partial r}{\partial z} = 1+1+1 = 3$$

$$\text{div}\,\hat{r} = \text{div}\,(\frac{\mathbf{r}}{r}) = \frac{\partial}{\partial x}(\frac{\mathbf{x}}{r}) + \frac{\partial}{\partial y} + (\frac{\mathbf{y}}{r}) + \frac{\partial}{\partial z}(\frac{\mathbf{z}}{r})$$

$$= \frac{\partial}{\partial x}(\frac{x}{(x^2+y^2+z^2)^{1/2}}) + \text{two similar terms}$$

$$= \frac{r.1 - x.\frac{1}{2}.\frac{1}{(x^2+y^2+z^2)^{1/2}}.2x}{(x^2+y^2+z^2)} + \text{two similar terms}$$

$$= \frac{r - \frac{x^2}{r}}{r^2} + \text{two similar terms} = (\frac{1}{r} + \frac{x^2}{r^3}) + \text{two similar terms}$$

$$= (\frac{1}{r} - \frac{x^2}{r^3}) + (\frac{1}{r} - \frac{x^2}{r^3}) + (\frac{1}{r} - \frac{z^2}{r^3}) = \frac{3}{r} - \frac{1}{r}(\frac{x^2+y^2+z^2}{r^2}) = \frac{3}{r} - \frac{1}{r} = \frac{2}{r}$$

Thus div $\hat{r} = \frac{2}{r}$

Ex.6 Show that div $(\frac{\hat{r}}{r^2})$ is zero

Sol. We have

$$\text{div}\,(\frac{\hat{r}}{r^2}) = (\frac{r}{r^3}) = \text{div}\,[\frac{\mathbf{i}x + \mathbf{j}y + \mathbf{k}z}{(x^2+y^2+z^2)^{3/2}}]$$

$$= \frac{\partial}{\partial x}\frac{x}{(x^2+y^2+z^2)^{3/2}} + \text{two similar terms}$$

$$= \frac{(x^2+y^2+z^2)^{\frac{3}{2}}.1 - x.\frac{3}{2}(x^2+y^2+z^2)^{\frac{1}{2}}.2x}{(x^2+y^2+z^2)^3} + \text{two similar terms}$$

$$= [\frac{1}{(x^2+y^2+z^2)^{3/2}} - \frac{3x^2}{(x^2+y^2+z^2)^{3/2}}] + \text{two similar terms}$$

$$= (\frac{1}{r^3} - \frac{3x^2}{r^5}) + \text{two similar terms}$$

$$= (\frac{1}{r^3} - \frac{3x^2}{r^5}) + (\frac{1}{r^3} - \frac{3y^2}{r^5}) + (\frac{1}{r^3} - \frac{3z^2}{r^5})$$

$$= \frac{3}{r^3} - \frac{3}{r^5}(x^2+y^2+z^2)\ \frac{3}{r^3} - \frac{3}{r^5}(r^2)$$

$$= \frac{3}{r^3} - \frac{3}{r^3} = 0$$

Ex.7 What must be the value of p so that the vector field **p** represented by $\mathbf{p} = \mathbf{i}(3x + py - 2z) + \mathbf{j}(2x + y - 2pz) + \mathbf{k}(4px - 3pz)$ is a solenoidal field ?

Sol. We want div p = 0

Now $\quad \text{div } p = \frac{\partial}{\partial x}p_x + \frac{\partial}{\partial y}p_y + \frac{\partial}{\partial z}p$

$$= \frac{\partial}{\partial x}(3x + py - 2z) + \frac{\partial}{\partial y}(2x + y - 2pz)$$

$$= 3+1-3 = 4 - 3p$$

$\therefore \quad \text{div } p = 0 \text{ if } \quad 4 - 3p = 0$

Or $\quad = 4/3$

Ex.8 Evaluate grad div B and div (grad u)

Sol. We have

$$\text{Grad (div B)} = \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right)$$

$$= \left[\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\mathbf{k}\right]\left[\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}\right]$$

$$= \mathbf{i}\left(\frac{\partial^2 B_x}{\partial x^2} + \frac{\partial^2 B_y}{\partial x\,\partial y} + \frac{\partial^2 B_z}{\partial x\,\partial z}\right) + \mathbf{j}\left(\frac{\partial^2 B_x}{\partial y\,\partial x} + \frac{\partial^2 B_y}{\partial y^2} + \frac{\partial^2 B_z}{\partial y\,\partial z}\right) + \mathbf{k}\left(\frac{\partial^2 B_x}{\partial z\,\partial x} + \frac{\partial^2 B_y}{\partial z\,\partial y} + \frac{\partial^2 B_z}{\partial z^2}\right)$$

Again,

$$\text{Div (grad u)} = \text{div}\left(\mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial z}{\partial z}\right)$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial u}{\partial z}\right)$$

$$= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)u$$

Ex.9 A central force field **F** is $\mathbf{F} = k\,r^n\mathbf{r}$. The field **F** is solenoidal such that div $\mathbf{F} = 0$.

Sol. Since $k \neq 0$; for field to be solenoidal, so

$$\text{div}(r^n r) = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\hat{k}\right).r^n\,(x\hat{\imath} + y\hat{\jmath} + z\hat{k})$$

$$= \frac{\partial}{\partial x}(r^n x) + \frac{\partial}{\partial y}(r^n y) + \frac{\partial}{\partial x}(r^n z)$$

Let us evaluate $\frac{\partial}{\partial x}(r^n x)$,

$$\frac{\partial}{\partial x}(r^n x) = \frac{\partial}{\partial x}\left[(x^2 + y^2 + z^2)^{\frac{n}{2}}.x\right]$$

$$= (x^2 + y^2 + z^2)^{\frac{n}{2}}.1 + \frac{n}{2}(x^2 + y^2 + z^2)^{\frac{n}{2}-1} \times 2x$$

$$= r^n + nx^2(r^2)^{\frac{n}{2}-1} = r^n + nx^2 r^{n-2}$$

Similarly, $\frac{\partial}{\partial x}(r^n y) = r^n + ny^2 r^{n-2}$

and $\quad \frac{\partial}{\partial x}(r^n z) = r^n + nz^2 r^{n-2}$

$$\text{div}(r^n r) = 3r^n + (x^2 + y^2 + z^2)r^{n-2}$$

$$= 3r^n + nr^n = (3 + n)r^n$$

In order that div $F = 0$; $3 + n = 0$

Or $\quad n = -3$

$F(r) = k/r^3$

Ex.10 Show that the position vector $\mathbf{r} = (\mathbf{i}x + \mathbf{j}y + \mathbf{k}z)$ is irrotational

Sol. We have to show that curl $\mathbf{r} = 0$

$$\text{Now curl } \mathbf{r} = \nabla \times \mathbf{r} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = \mathbf{i}\left(\frac{\partial z}{\partial y} - \frac{\partial y}{\partial z}\right) + \text{two similar terms}$$

Now as z does not depend on x and y and so on, we have each of $\frac{\partial z}{\partial y}$ and $\frac{\partial y}{\partial z}$ equal to zero.

Hence $\quad$ curl $\mathbf{r} = \mathbf{i}(0) + \mathbf{j}(0) + \mathbf{k}(0) = 0$

Since curl **r** is zero, **r** is an irrotational vector.

Ex.11 Given that the vector **E** defined by.

$\mathbf{E} = (2x - 5y + pz)\mathbf{i} + (qx + 3y)\mathbf{j} + (ry + 6z + 3x)\mathbf{k}$ is irrotational, find p, q, r.

Sol. We have

$$(\text{curl E})_x = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) \text{and so on.}$$

Hence $(\text{curl E})_x = (r - 0) = 0$

$(\text{curl E})_y = (p - 3) = 0$

$(\text{curl E})_z = (q - 5) = 0$

Thus, $\quad p = 0, q = 5, r = 0$

Ex.12 If $\mathbf{B} = yz^2\mathbf{i} - xyz\mathbf{j} + 3x^2z^3\ \mathbf{k}$, find curl **B** at the point (0, 1, 1).

Sol. We have

$$(\text{curl B})_x = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) = \frac{\partial}{\partial y}(3x^2z^3) - \frac{\partial}{\partial y}(-xyz) = 0 + xy = xy$$

Similarly, $\quad (\text{curl B})_y = \frac{\partial}{\partial y}(yz^2) - \frac{\partial}{\partial y}(3x^2z^3) = 2yz + 6xz^3$

$$(\text{curl B})_z = \frac{\partial}{\partial y}(-xyz) - \frac{\partial}{\partial y}(yz^2) = -yz - z^2$$

$\therefore$ curl B $= \mathbf{i}(xy) + \mathbf{j}(2yz + 6xz^3) - \mathbf{k}(yz + z^2)$

Thus $(\text{curl }\mathbf{B})_{0,1,1} = \mathbf{i}(0 \times 1) + \mathbf{j}(2 \times 1 \times 1 + 6 \times 0 \times 1^3) - \mathbf{k}(1 \times 1 \times 1^2)$

$= 2(\mathbf{j} - \mathbf{k})$

Ex.13 If $\mathbf{F} = (xy)\mathbf{i} - (yz)\mathbf{j} + (zx)\mathbf{k}$, find curl **F** and curl(curl **F**). Evaluate it at (0, 0, -1).

Sol. We have

$$\text{Curl }\mathbf{F} = \mathbf{i}\left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right) + \mathbf{k}\left(\frac{\partial F_y}{\partial x} - \frac{\partial x}{\partial y}\right)$$

$$= \mathbf{i}\left[\frac{\partial}{\partial y}(zx) - \frac{\partial}{\partial z}(-yz)\right] + \mathbf{j}\left[\frac{\partial}{\partial z}(xy) - \frac{\partial}{\partial x}(yz)\right] + \mathbf{k}\left[\frac{\partial}{\partial x}(-yz) - \frac{\partial}{\partial y}(xy)\right]$$

$$= \mathbf{i}(+y) + \mathbf{j}(-z) + \mathbf{k}(-x)$$

$$\therefore \text{Curl (curl }\mathbf{F}) = \mathbf{i}\left[\frac{\partial}{\partial y}(-x) - \mathbf{j}\frac{\partial}{\partial z}(-z)\right] + \mathbf{j}\left[\frac{\partial}{\partial z}(y) - \frac{\partial}{\partial x}(x)\right] + \mathbf{k}\left[\frac{\partial}{\partial x}(-z) - \frac{\partial}{\partial y}(+y)\right]$$

$$= \mathbf{i}(+1) + \mathbf{j}(+1) + \mathbf{k}(-1)$$

Curl(curl **F**) $= \mathbf{i} + \mathbf{j} + \mathbf{k}$

It has the same value at all points and value at (0, 0, −1) is also $\mathbf{i} + \mathbf{j} + \mathbf{k}$.

Based on Equation of continuity, Displacement current

Ex.14 A parallel plate capacitor having circular plates of radius 5 cm is being charged. If the electric field between the plates during charging is changing at the rate of 10^{12} V/ms, find the displacement current between the plates.

Sol. The displacement current is

$$I_d = \varepsilon_0 \frac{d\emptyset_E}{dt}$$

$$= \frac{d}{dt}(\pi R^2 E)$$

$$= \varepsilon_0 \pi R^2 \frac{dE}{dt}$$

$$k = \alpha = {}^{2\pi}/_{\lambda} = \frac{\omega}{c}$$
$$= \frac{10^8}{3 \times 10^8} = \frac{1}{3} m^{-1}$$

The frequency of the wave is

$$\nu = \frac{\omega}{2\pi} = \frac{10^8}{2 \times 3.14} = 1.67 \times 10^7 Hz$$

The magnetic field is similar in form to the electric field:

$\mathbf{B} = B_0 \exp[-\mathbf{i}\,(10^8\, t + \frac{1}{3} z)]$

The magnitude of B_0 is given by

$$B = \frac{E_0}{c} = \frac{60}{c}$$
$$= \frac{60}{3 \times 10^8} = 2 \times 10^{-7} \text{tesla}$$

Ex.21 A plane electromagnetic wave is travelling in the $-\hat{x}$ direction. Its frequency is 100 MHz and the electric field is perpendicular to $\hat{z}$ direction. Write down the expressions for the **E** and **B** fields that specify the wave.

Sol. Since the wave is travelling in the $-\hat{x}$ direction, the **E** field is normal to $\hat{x}$. Also, **E** is perpendicular to $\hat{z}$ direction. Therefore, **E** is in the $\hat{y}$ direction. The expression for the **E** field can be written as

$$\mathbf{E} = E_0 \hat{y} \cos(kx + \omega t)$$

Where $\omega = 2\pi\nu = 2\pi \times 10^8$ Hz

And $k = \frac{\omega}{c} = \frac{2\pi}{3} m^{-1}$

Thus,

$$\mathbf{E} = E_0 \hat{y} \cos[2\pi(\frac{x}{3} x + 10^8 t)$$

The corresponding **B** field is given by

$$B = \frac{1}{\omega}(k \times E) = \frac{k}{\omega}(\hat{k} \times E)$$
$$= \frac{1}{c}\hat{k} \times E$$
$$= \frac{1}{c}(-\hat{x} \times \hat{y}) E_0 \cos[2\pi\left(\frac{x}{3} x + 10^8 t\right)]$$
$$= -\frac{E_0}{c}\hat{z} \cos[2\pi\left(\frac{x}{3} x + 10^8 t\right)]$$

Ex.22 The electric field of an electromagnetic wave in free space is given by

$E_x = 0, E_y = 50\sin(2\pi \times 10^8 t - \frac{2\pi}{5} x), E_z = 0$

Where all the quantities are in SI units. Determine a) the wavelength of the wave b) the direction of propagation of the wave and c) the direction of the magnetic field.

Sol. The problem can be solved by comparing the given expression for E_y with the standard expression

$\mathbf{E_y} = E_0 \sin(\omega t - \mathbf{k}x)$,

$\omega = 2\pi \times 10^8$ rad/s, $\mathbf{k} = \frac{2\pi}{5} m^{-1}$

a) Wavelength $\lambda = \frac{2\pi}{k} = \frac{2\pi}{\frac{2\pi/5}{5}} = 5m$

b) The direction of propagation is the +x direction.

c) Since **E** is the +y direction and **k** is along the +x direction, the **B** field is along the + z direction.

Based on electromagnetic wave travelling in a dielectric medium

Ex.23 A plane electromagnetic wave travelling in positive z-direction in an unbounded lossless dielectric medium with relative permeability $\mu_r = 1$ and relative permittivity $\varepsilon_r = 3$ has a peak electric field intensity $E_0 = 6\ \text{V/m}$. Find

i) The speed of the wave

ii) the intrinsic impedance of the medium

iii) The peak magnetic field intensity (H_0), and

iv) The peak Poynting vector S (z, t).

Sol. $E_0 = \sqrt{\left(E^2{}_{ox} + E^2{}_{oy}\right)} = 6\text{V/m},\ \varepsilon_r = 3,\ \mu_r = 1.$

i) The speed of electromagnetic wave

$$v = \frac{1}{\sqrt{\mu\varepsilon}} = \frac{1}{\sqrt{(\mu_r\mu_0\varepsilon_r\varepsilon_0)}}$$
$$= \frac{1}{\sqrt{(\mu_0\varepsilon_0)}} \cdot \frac{1}{\sqrt{(\mu_r\varepsilon_r)}} = \frac{c}{\sqrt{(\mu_r\varepsilon_r)}}$$
$$= \frac{3\times10^8}{\sqrt{(1\times3)}} = 1\cdot73\times10^8\text{m/s}.$$

ii) Impedance of medium

$$z = \sqrt{\left(\frac{\mu}{\varepsilon}\right)} = \sqrt{\left(\frac{\mu_r\mu_0}{\varepsilon_r\varepsilon_0}\right)} = \sqrt{\left(\frac{\mu_0}{\varepsilon_0}\right)} \cdot \sqrt{\left(\frac{\mu_r}{\varepsilon_r}\right)}$$
$$= \sqrt{\left(\frac{4\pi\times10^{-7}}{8\cdot86\times10^{-12}}\right)} \cdot \sqrt{\left(\frac{1}{3}\right)} = \frac{376\cdot6}{\sqrt{3}}$$
$$= 217\cdot6\ \Omega.$$

iii) Peak value of magnetic field

$$H_0 = \frac{E_0}{Z} = \frac{6}{217\cdot6} = 2\cdot76\times10^{-2}\text{A/m}.$$

iv) Poynting vector $\quad S = E \times H$

Peak Poynting vector $= E_0H_0 = \dfrac{{E_0}^2}{Z}$

$$= \frac{6^2}{217\cdot6} = 0\cdot165\ \text{W/m}^2$$

Ex.24 In a homogenous non-conducting medium the electric and magnetic fields of electromagnetic wave are given by:

$E = 30\pi\hat{z}\exp\left[\mathbf{i}\left(\omega t - \frac{4}{3}y\right)\right]$ V/m and $H = 1\hat{x}\exp\left[\mathbf{i}\left(\omega t - \frac{4}{3}y\right)\right]$ A/m

It is given that for the medium, $\mu_r = 1$. Calculate a) ε_r b) the velocity of light in this medium c) ω

Sol. a) From the equation:

$$Z = \frac{E}{H} = 377\sqrt{\frac{\mu_r}{\varepsilon_r}}$$

$$30\pi = 377\sqrt{\frac{1}{\varepsilon_r}}$$

Or $\varepsilon_r = (\frac{377}{30\pi})^2 = 16$

b) The velocity of light in the medium is

$$v = \frac{1}{\sqrt{\mu\varepsilon}} = \frac{1}{\sqrt{\mu_0\mu_r\varepsilon_0\varepsilon_r}}$$
$$= (\frac{1}{\sqrt{\mu_0\varepsilon_0}})(\frac{1}{\sqrt{\mu_r\varepsilon_r}})$$
$$= \frac{c}{\sqrt{\mu_r\varepsilon_r}} = \frac{c}{\sqrt{1\times 16}}$$
$$= \frac{c}{4} = \frac{3\times 10^8}{4}$$
$$= 7.5\times 10^7 \mathrm{m/s}$$

c) $\omega = vk = \left(\frac{c}{4}\right)\left(\frac{4}{3}\right) = \frac{c}{3} = 10^8 \mathrm{rad/s}$

Based on skin depth

Ex.25. Calculate the skin depth for 3 MHz electromagnetic wave through copper. (Given conductivity $\sigma = 6\times 10^7 \mathrm{mho/m}$, $\mu = 4\pi\times 10^{-7}$ Henry/m.

Sol. $\delta = \sqrt{\frac{2}{\mu\sigma\omega}} = \sqrt{\frac{2}{\mu\sigma(2\pi v)}} = \sqrt{\frac{2}{4\pi\times 10^{-7}\times 6\times 10^7(2\pi\times 3\times 10^6)}} = 37.5\mu\mathrm{m}$

Ex.26 The constitution parameter of aluminium is given by $\mu_r = 1$, $\varepsilon_r = 1$ and $\sigma = 3.54\times 10^7 \mathrm{mho/m}$. Find the frequency for which the skin depth/penetration depth of aluminium is 0.01mm.

Sol. $\delta = \sqrt{\frac{2}{\mu\sigma\omega}} = \sqrt{\frac{1}{\mu\sigma(\pi v)}}$

Or $v = \sqrt{\frac{1}{\pi\delta^2\mu\sigma}}$

$$= \frac{1}{3.14\times(0.01\times 10^{-3})^2\times 1\times 3.54\times 10^7}$$
$$= \frac{1000}{3.14\times 3.54} = 89.96 \text{ Hz}$$

Ex.27 For silver, $\sigma = 5.0 \mathrm{MS/m}$. At what frequency will the depth of penetration δ be 1mm?

Sol. $v = \sqrt{\frac{1}{\pi\delta^2\mu\sigma}} = 84.4\text{kHz}$

Ex.28 Show that for a good conductor the magnetic field lags the electric field by 45^0. Determine the ratio of their amplitudes.

Sol. $\alpha + i\beta = (1+i)\sqrt{\frac{\mu\sigma\omega}{2}}$

$\alpha = \sqrt{\frac{\mu\sigma\omega}{2}}$ and $\beta = \sqrt{\frac{\mu\sigma\omega}{2}}$.

$\emptyset = \tan^{-1}\frac{\beta}{\alpha} = 45^0$

*****Review Questions*****

Based on Simple harmonic motion

1.What is meant by a harmonic oscillator? Obtain expressions for (1) its displacement and velocity at a given instant (2) time period and frequency

2.Show that for a simple harmonic oscillator, mechanical energy remains conserved and that its energy is on an average, half kinetic and half potential in form. At what particular displacement is this exactly so? What is the ratio between its kinetic and potential energies at a displacement equal to half of its amplitude?

3. Solve the differential equation $\frac{d^2x}{dt^2} + \omega^2 x = 0$ to obtain the expression $x = A\sin(\omega t + \delta)$ for the displacement of a particle executing S.H.M.

4.Show that for a S.H.M may be expressed as either a sine or a cosine wave function, there being only a difference of initial phase in the two cases.

5. Is it really possible to construct a truly simple pendulum? What are the drawbacks of simple pendulum?

6.How does a compound pendulum differ from a simple pendulum? Obtain an expression for its time period and mention its points of superiority over simple pendulum.

7. Define centres of suspension and oscillation of a compound pendulum and show that they are interchangeable. What length of the pendulum has its minimum time period?

Based on Equation of continuity, displacement current

8.State the equation of continuity.

9.What is the physical significance of equation of continuity?

10.What is meant by displacement current?

11.In what way the displacement current is different from the conventional current?

12.What is the value of displacement current when a capacitor becomes fully charged? Give explicit reasoning for your answer.

13.Give the equation of continuity of electromagnetic theory. Explain the inconsistency of Ampere's law for transient currents. How was the law modified in its generalized form to overcome the inconsistency?

14.Discuss in brief the inconsistencies in Ampere's law and describe how Maxwell fixed up this. Further discuss in brief the characteristics of displacement current.

15.Illustrate mathematically how and under what conditions does Ampere's circuital law fail. How did Maxwell modify Ampere's law to make it consistent under all conditions? Give the mathematical justification to prove this consistency

Based on Maxwell's equation

16.Write Maxwell's equation differential form

17.Write Maxwell's equation in integral form.

18.One of the Maxwell's equations is $\nabla \cdot B = 0$. What is its physical implication?

19.What was the most important consequence of Maxwell's equations.

20.Explain how Maxwell's equations were developed. Write them in integral and differential forms.

21.Write Maxwell's equation in integral form. Discuss the physical meaning of each of these.

22.Enumerate Maxwell's equations and show how they predict the existence of electromagnetic waves.

Based on electromagnetic waves in free space

23.What is the velocity of electromagnetic waves in free space?

24.Show that the electromagnetic waves are transverse waves.

25.Show that in an electromagnetic wave the E and B fields are perpendicular to each other and also perpendicular to the direction of propagation.

26.What is wave impedance? What is its value for free space?
27.Write the expression for the refractive index of a medium in terms of electric and magnetic quantities.
28.State Maxwell's equations for the electromagnetic fields and obtain the wave equations for E and B in free space.
29.Discuss the propagation of plane electromagnetic waves through free space. Establish the transverse nature of these waves. What is the expression for wave impedance?

Based on Poynting theorem and Poynting vector
30.State the Poynting theorem. Explain the term Poynting vector.
31.What is Poynting vector and what does it represent?
32.State Poynting theorem.
33.What is Poynting vector? How is the Poynting theorem derived from Maxwell's curl equations? Explain Poynting theorem.

Based on em waves in isotropic dielectric
34.Write down Maxwell's equations for electromagnetic fields in a homogeneous isotropic dielectric. Solve these equations to get the velocity of propagation of electromagnetic waves. Why do we regard these waves as transverse? Show that the wave energy is equally shared between the electric and magnetic fields.
35.Discuss the propagation of plane electromagnetic waves in a non-conducting medium.
36.Obtain the wave equation for a plane E M wave in an isotropic dielectric medium and show that its velocity of propagation is less than the speed of light.
37.(a) Write Maxwell's equations in differential form and give their significance.
(b) Derive the wave equation for an isotropic dielectric medium. Prove the orthogonality of E, H and K vectors. Find the wave impedance of the medium.

Based on Wave in conducting medium
38.What do you understand by skin depth?
39.(a) Discuss the propagation of monochromatic plane electromagnetic waves in a conducting medium. (b) Show that in a good conductor the magnetic field lags the electric field by 45°. Determine the ratio of their amplitudes.
40.Define skin depth. Show that in case of good conductor, the skin depth is given by $\delta = (\frac{2}{\omega\sigma\mu})^{1/2}$
Show that inside the conducting medium electromagnetic wave is damped and obtain an expression for skin depth.
41.Obtain the equation of plane em wave in a conducting medium. Prove the orthogonality of electric vector E, intensity of magnetic field H and propagation wave vector $\widehat{k}$.

****

Unit - III

Interference: Interference by division of wave front (Young's Doble slit experiment, Fresnel biprism), Interference by division of amplitude (Thin film, Newton's rings, Michelson interferometer), Coherence and coherent sources.

Chapter 3

Interference of Light

Introduction

"The phenomena of modification of amplitude/intensity of light at point, illuminated by two or more than two light waves, is termed as interference of light"

Here we have used the term light waves in place of light rays due to obvious reasons. Some of the phenomena of light could not be explained on the basis of ray optics. Interference of light is one of them.

The literal meaning of interference of light is the illumination of a point with several waves of light. These individual waves of light may differ in several physical properties, i.e. velocity, frequency, wavelength, amplitude, phase etc. Though the interference of light will take place at that point with these waves but the pattern of intensity obtained will be such haphazard, so that we cannot make any measurements or we cannot utilize the outcomes of such phenomena for any fruitful purpose, which is the ultimate objective of studying any physical phenomena. Thus, to obtain well defined and sustained intensity pattern we need the interfering waves to have some physical condition. So, what are those conditions, how we obtain them, how the mathematical expressions are developed for the ultimate measurements in the interference phenomena, is our matter of discussion in the current chapter.

As of now we all understand that light is an electromagnetic wave having electric and magnetic field vectors, oscillating perpendicular to the propagation of light as well as perpendicular to each other. Also, the intensity of light is decided by the amplitude of either the electric field or the magnetic field. Thus, the modification of intensity of light at any point due to interference of light can be understood by studying the superposition of two or more electric or magnetic vectors of light at that point.
As the intensity can be defined by either electric or magnetic vector so here, we will treat the electric field vector to understand the interference of light.

3.1 Interference of Light Waves

Superposition of two or more than two light waves is the mathematical inference of interference. Let us now understand superposition of two electric field vectors to understand the interference of light. And finally, we will derive various mathematical expressions for doing further measurements for various purposes.

3.1.1 Superposition of Two Light Waves

Let there are two light wave trains whose electric vectors are oscillating with same frequency. Obviously, the respective magnetic field vectors will be oscillating with the same frequency and is perpendicular to their respective electric fields.

The progressive wave equations of oscillating electric field for a light beam can be written as:

$$E_y = E_{y0} \sin(kx + \omega t) \qquad (1)$$

And the corresponding magnetic field vector be written as:

$$B_z = B_{z0} \sin(kx + \omega t + \phi) \qquad (2)$$

As seen from equations (1) and (2), the electric field and magnetic field vector are oscillating in y-axis and z-axis respectively and the wave is progressing along x-axis. The intensity of light can be determined by either $\frac{1}{2}\,\varepsilon\,E^2$ or $\frac{1}{2}\,\mu\,B^2$. For the interference purpose and for simplicity we can take either of the two perpendicularly oscillating vector and can also omit the symbols representing the direction of oscillation and direction of propagation.

3.1.2 Mathematical Treatment of Superposition of Two Light Waves

Keeping the above said points into our mind, let there be two light beams 1 and 2 originating from two identical sources and their resultant, as shown in Fig.3.1.

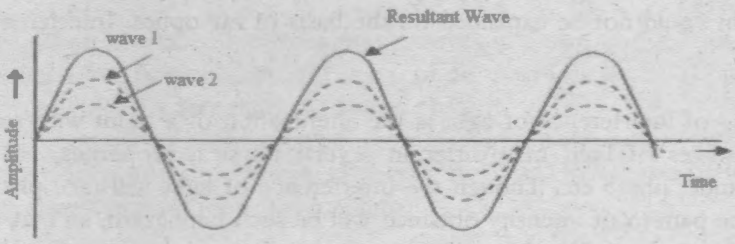


Fig. 3.1 Superposition of two waves

The equations of electric field vectors E_1 and E_2 of the two wave trains of the individual source at a point on the screen can be written as;

$$E_1 = E_{01} \sin \omega t \qquad (3)$$

$$E_2 = E_{02} \sin(\omega t + \phi) \qquad (4)$$

Where E_1 and E_2 are the instantaneous amplitudes at the point of observation on the screen. E_{01} and E_{02} are the electric fields amplitudes of each wave and ϕ is the phase difference between two light waves, produced due to different paths travelled by the two waves. The phase difference ϕ, due to the path difference, Δ, between the two waves, is given according to the relation as:

Phase difference, $\phi = \frac{2\pi}{\lambda}$ (path difference)

$$\phi = \frac{2\pi}{\lambda}\Delta, \qquad (\lambda \text{ being the wavelength of the waves.})$$

However, the point on the screen is illuminated by the intensities of the two light waves which may or may not be in phase. The question that comes into our mind is that: what should be the resultant intensity of light at that point? This means that we have to find out the resultant of two oscillating electric fields giving the intensity of light at that point.

To find the resultant intensity at any point, which is proportional to E^2, we have to find out the resultant of two electric field vectors, given by the principle of superposition.

$$E_R = E_1 + E_2$$
$$E_R = E_{01} \sin \omega t + E_{02} \sin(\omega t + \phi)$$
$$= E_{01} \sin \omega t + E_{02} (\sin \omega t \cos\phi + \cos\omega t \sin\phi)$$
$$= (E_{01} + E_{02} \cos \phi)\sin \omega t + E_{02} \sin \phi \cos \omega t$$
$$= E_0 \cos\delta \sin \omega t + E_0 \sin \delta \cos \omega t$$
$$E_R = E_0 \sin (\omega t + \delta) \quad (5)$$

$$\text{where } E_{01} + E_{02} \cos \phi = E_0 \cos\delta \quad (6)$$
$$E_{02} \sin\phi = E_0 \sin\delta \quad (7)$$

Here, E_R is the instantaneous amplitude of resultant electric field vector in the electromagnetic field oscillating with same frequency ω, at the point on the screen where two light waves having instantaneous amplitude given by (3) and (4), superimpose and the amplitude of the resultant wave E_0 can be calculated by squaring and adding (6) and (7), as:

$$E_0 = \sqrt{E_{01}^2 + E_{02}^2 + 2E_{01} E_{02} \cos\phi} \quad (8)$$

And the final phase angle δ can be found out by dividing (7) by (6). As:

$$\tan \delta = \frac{E_{02} \sin\phi}{E_{01} + E_{02} \cos \phi} \quad (9)$$

Let us analyze expression (5) to find the intensity at the point of superposition. The intensity of light at a point is proportional to E_R^2 so more appropriately it is calculated as the average over a complete cycle of the oscillating electric field i.e.

$$I = \frac{1}{2}\varepsilon < E_R^2 >$$
$$I = \frac{1}{2}\varepsilon < E_0{}^2 \sin^2 (\omega t + \delta) >$$

Since the average of $\sin^2$ over a cycle gives $\frac{1}{2}$. So,

$$I = \frac{1}{4}\varepsilon E_0{}^2$$

Thus, the intensity at the point of interference will be decided by the amplitude of the resultant wave E_0.

From (8), it is clear that E_0 is a function of ϕ i.e., phase difference between the two interfering waves. If the phase difference is fluctuating i.e., if ϕ is a function of time, intensity will vary at the same point with time. This means for sustainability of the intensity at a point, ϕ should not vary with time at the same place.

The phase difference between the two waves arises mainly due to the following two reasons;

(i) If the frequencies of the oscillation of the electric fields of the light waves are different then phase will develop between them and the difference in the phase will be fluctuating.

(ii) If the initial phase difference between the two waves is kept constant by keeping the frequencies of the two sources same even then the different phase difference between the two waves will be seen at the different points in the interfering region, which is produced due to different paths traversed by the two waves.

This shows that the intensity variation at different points in the interfering region is due to the phase difference ϕ, between two waves. If the phase difference between the two light waves is constant, the resultant amplitude at the point will be constant, otherwise, intensity will change with time, i.e., if $\cos\phi$ is function of time, the amplitude/ intensity at the point will vary with time. This implies that for the sustainability of the intensity (pattern) at the point, the phase difference must be constant. At different points on the screen, the path difference/ phase difference between the two superimposing waves will be different and so the intensity of light will be different at different points on the screen. This variation in the intensity on the screen due to superposition of the two waves is given a term interference of light. So, the intensity of light on the screen is a function of $\cos\phi$, which is constant at a point if ϕ is constant but varies at different points on the screen due to different path difference. We can find the points of varying intensity with the $\cos\phi$ factor e.g., the point is of maximum intensity if $\cos\phi$ is 1 minimum intensity, if $\cos\phi$ is -1 and average intensity if $\cos\phi$ is 0. As the intensity at arbitrary point on the screen is given by:

$$I = E_{01}^2 + E_{02}^2 + 2E_{01}\,E_{02}\cos\phi$$

$$I = I_1+I_2 +\sqrt{I_1 I_2}\,2\cos\phi \qquad (10)$$

Points of Maximum Intensity

$\cos\phi = 1$ for I to be maximum

The phase difference between two interfering waves must be thus an even multiple of π i.e. Phase difference, $\phi = 2n\,\pi$ or path difference must be even multiple of $\lambda/2$, i.e.

Path difference $\nabla = 2n\,\lambda/2$

$$I_{Max} = E_{01}^2 + E_{02}^2 + 2E_{01}\,E_{02} \qquad (11)$$

$$I_{Max} = I_1 + I_2 + 2\sqrt{I_1 I_2} \qquad (12)$$

Points of Minimum Intensity

$\cos\phi = -1$ for I to be minimum

The phase difference between two interfering waves must be thus an odd multiple of π i.e., Phase difference, $\phi = (2n \pm 1)\,\pi$ or path difference must be even multiple of $\lambda/2$, i.e.

Path difference $\nabla = (2n \pm 1)\,\lambda/2$

where, n is the order of the points of maxima or minima.

The minimum intensity in the interference pattern on the screen comes out to be;

$$I_{Min} = E_{01}^2 + E_{02}^2 - 2E_{01}\,E_{02} \qquad (13)$$

$$I_{Min} = I_1 + I_2 - 2\sqrt{I_1 I_2} \qquad (14)$$

Points of Average Intensity

The points of average intensity will be between the points of maxima and minima.

$\cos\phi = 0$, for I to be average of maximum and minimum intensity.

The phase difference between two interfering waves must thus be a multiple of $\pi/2$.

Phase difference, $\phi = n\pi/2$ or path difference must be multiple of $\lambda/4$, i.e.

Path difference $\nabla = n\lambda\,/2$

The average intensity can also be calculated as;

$$I_{avg} = \frac{\int_0^{2\pi} I\, d\phi}{\int_0^{2\pi} d\phi} \qquad \text{(Average intensity over a complete cycle)}$$

$$= \frac{\int_0^{2\pi} (E_{01}^2 + E_{02}^2 + 2E_{01}\, E_{02} \cos\phi)\, d\phi}{\int_0^{2\pi} d\phi}$$

$$= \frac{(E_{01}^2 + E_{02}^2)2\pi}{2\pi}$$

$$= E_{01}^2 + E_{02}^2 \qquad (15)$$

$$= I_1 + I_2 \qquad (16)$$

From the equations (11), (13) and (15), it is clearly seen that $2E_{01}\, E_{02}$, is the amount of intensity which is redistributed from minima to maxima and vice a versa. This treatment explains the concept of conservation of energy in the interference phenomena too. One can visualize the variation of intensity at different points on the screen due to redistribution of energy on account of superposition of two light wave trains.

Intensity Curve of Interference of Two Light Waves

If the intensities of both the waves is same at every time, then equation (10) takes the form:

$I_1 = I_2 = I_0$ (let) then

$I = 2I_0 + I_0\, 2\cos\phi$

Or $I = 2I_0\ (1 + \cos\phi)$

Or $I = 2\, I_0\ .\ 2\cos^2 \frac{\phi}{2}$

Or $I = 4I_0 \cos^2 \frac{\phi}{2}$ (17)

Thus, the maximum, minimum and the average value of intensity in the pattern will be given as;

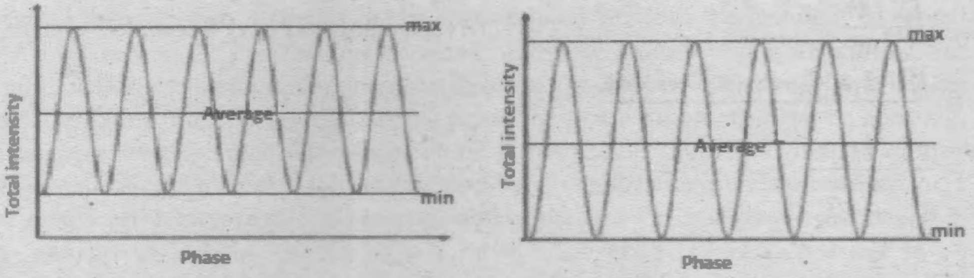
$I_{Max.} = 4I_0$ (18)

$I_{Min.} = 0$ (19)

$I_{avg} = 2\, I_0$ (20)

It means for a good contrast on the screen the amplitude/intensity of two interfering light waves should be same or nearly same at least.

From the above mathematical treatment of two light waves, it can be concluded that the interference of two waves shows the intensity variation in the interfering region. Also, an observable, well-defined and sustained interference pattern can be seen on the screen under certain conditions only. According to equation (10) and (11), the intensity variation curves on the screen will be given as Fig.3.2 (a) and (b) respectively.



(a) Intensity variation curve for $I_1 \neq I_2$ **(b) Intensity variation curve for $I_1 = I_2 = I_0$**

Fig. 3.2 YDSE Intensity Curve

3.2 Conditions for Observable, Well-defined and Sustained Interference Pattern

1. Observable : Highly coherent sources of light
2. Sustained : Phase difference must be constant at a point of interference
3. Well-defined : Amplitude/ Intensity of two waves should be nearly same

For good contrast, the experiment must be performed in a dark room with a black screen.
The use of the phenomena of interference of light for various purposes and applications for human kind can be understood only and only if one can study the interference which is sustained, well defined and have good contrast between points of maxima and minima. Else, interference of incoherent sources is of no use.

The next question now that comes into our mind is to how to get the perfect conditions in order to observe the required interference pattern?
The very first condition which is an important requisite also is: the presence of two coherent sources of light. But before this, we need to understand what coherency actually means. Coherent waves mean the oscillations of electric field vectors of two light waves are of same frequency. The amplitude of both the waves may be different, but it should vary with same rate i.e., phase between the waves should always be constant. First, we have to understand why two independent coherent sources of light are impossible to produce two independent coherent sources. Let us see how two sources of light waves of same frequency and constant phase difference can be produced which is the primary requirement to produce sustainable and observable interference phenomena. Also, our area of concern would be to know the limitations of these sources to remain coherent.

3.3 Coherent Sources

Superposition or interference takes place whenever the two sources are coherent or they are not. In case of sound waves, we get beats if there is a difference between the frequencies of the two superimposing sound waves and a slight difference in the form of beats can be detected by the human ear. Similarly, in case of light waves the superposition or interference will always take place if the two sources are coherent or not but the interference pattern will not be recognized due to human eye persistency and we cannot see sustained interference pattern if the sources we are using are not coherent. Hence, measurements can't be carried out, for its use for various applications. The study of any physical phenomena is done for its ultimate applications, and for this you need to carry out various measurements which further needs a sustained interference pattern. Therefore, we need two coherent sources for the interference of light waves.

By this time, we can learn that if we have to apply the phenomena of interference of light for various applications for human kind, we have to study the interference which is sustained, well defined and have good contrast between points of maxima and minima. Else, interference of incoherent sources is of no use.

The question now that arises is to how to get the required prefect conditions for the required interference pattern.

The foremost condition is to obtain two coherent sources of light. But before that, we get to know what coherency means. Coherent wave means the oscillations of electric field vectors of two light waves are of same frequency. The amplitude of both the waves may be different, but it should vary with same rate i.e. phase between the waves should always be constant. Before we proceed we have to understand why two independent coherent sources of light are impossible to produce sustained interference and how two sources of light waves of same frequency and constant phase difference can be generated which becomes the primary requirement to produce sustainable and observable interference phenomena.

Why two independent sources of light are not possible, to produce?

The electromagnetic waves can be produced by the oscillations of charges. Hertz's experiment and LC oscillations are the examples of this. But they are of very low frequencies comparable to the frequencies of light waves. The frequencies of visible spectrum can only be produced by excitation of atoms of lower energy levels to higher energy levels where they can stay for average time of nearly 10^{-8} s and after which they come back to the lower energy state. The frequency of the e.m radiation emitted by this de-excitation process is decided by the difference between these two levels. Here, we have to understand few facts about the radiation coming out of this process.

(1) It has been proved that the energy levels are not lines rather they are energy bands, so radiations coming out from the excitation and de-excitation process do not have a single frequency but a band of frequencies.

(2) It is a random process as it depends on how many atoms will get excited to reach to the higher energy level each time and also the average time of stay in the excited state may be different for different set of numbers of atoms. Therefore, it is uncertain that in each process how many will get excited and how many will be de-excited in the average time of 10^{-8} s. Although the supply of energy for exciting the lower energy level atoms is continuous but a bunch of different number of atoms absorb energy to go to higher energy level each time and also it is not necessary that this whole bunch stay in the excited state for a fixed average time interval to come again to the lower level.

3.4 Methods to Produce Two Coherent Sources

There are mainly two methods to generate two coherent sources which fulfill the conditions for producing sustained and well-defined interference pattern. This can further be utilized for developing the necessary formulations for interference phenomenon to use them for practical purposes.

a) By dividing a single wave-front: According to Huygens's theory of wave front, a wave front consists of the points which are in same phase within the limits of spatial coherence. So, by taking two parts of a wave front and making them to interfere, we can serve our purpose of two sources. Initially, it was done with the help of two prisms. As the prism deviates the wave front, towards its base, so by putting two prism joined at their bases, a wave front will be divided in to two parts in such a way that they can interfere, as shown in the Fig.3.3. The above principle can be accomplished by using Fresnel bi-prism, Fresnel's double mirror, Lloyd's single mirror (Fig.3.4) etc.

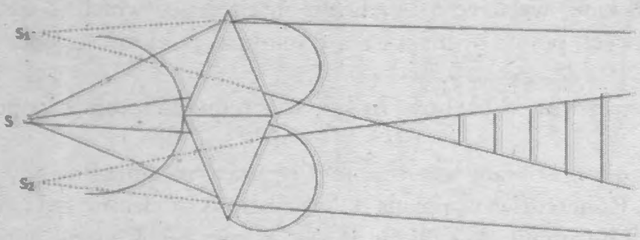


Fig.3.3. Division of single wave-front into two by a set of two prisms

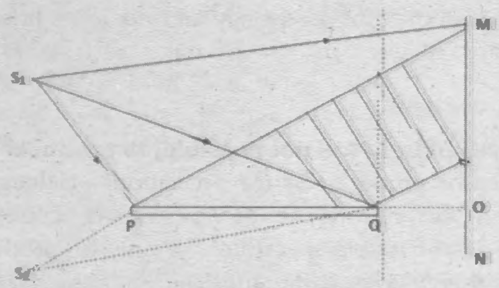


Fig.3.4. Lloyd's Single mirror

b) By dividing the amplitude: In this process, the amplitude of a single wave train is divided into two parts. As we know, the reflected and refracted waves of a same wave train have exactly the same frequency as of the original wave so if by any means any two waves out of these three waves (incident, reflected and refracted) are allowed to meet at any point, they will be treated as coming out from two independent coherent sources. They will fulfill the requirement for required interference phenomena. In general, this can be achieved by thins films. Any transparent film of appropriate thickness can serve the purpose for producing two coherent sources by division of amplitude as shown in Fig. 3.5. Brilliance of colours through soap bubbles, a compact disc, oil spread on water surface etc. are the examples of interference of light by division of amplitude using thin films.

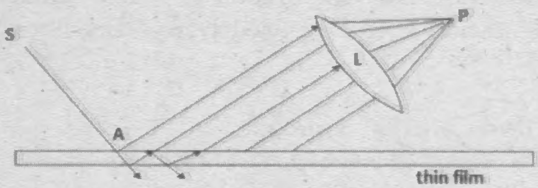


Fig.3.5. Interference by division of amplitude in thin films

3.5 Fresnel Biprism

Fresnel made a device that looked as if two prisms are joined base to base but with a single piece of glass known as Bi-Prism. This device is widely being used to generate two coherent sources by the division of wave front as shown in Fig.3.6. Nowadays, this is widely being used to produce two coherent sources Fig. 3.6. A biprism is actually is a single prism with an obtuse angle of 179° and the remaining two acute angles of 30′ each ($\alpha = 30'$)

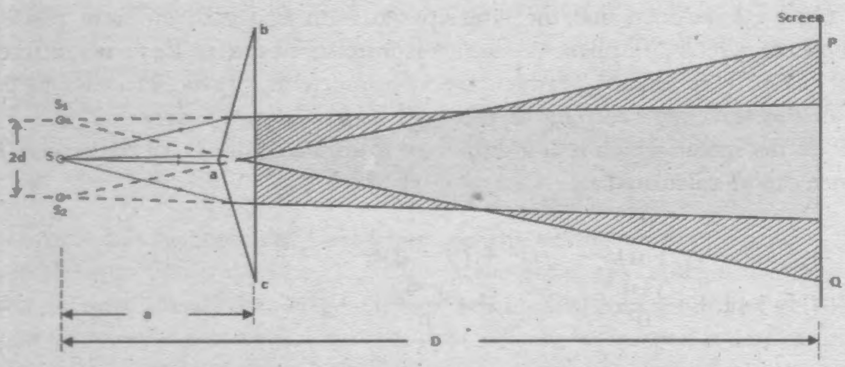


Fig. 3.6. Fresnel's biprism to produce two virtual sources

Here, S is a narrow vertical slit (perpendicular to the plane of paper), illuminated by a monochromatic source of light. The light from the slit S is then made to fall symmetrically on the biprism abc with its vertical refracting edge parallel to the light source S. When light falls on the upper half of the biprism, it bends downwards and appears to come from S_1 and if it falls on the lower half, it bends upwards as if coming from S_2. Each half thus produces a virtual image. These virtual images S_1 and S_2 act as two coherent sources. The cones of light PS_1Q and QS_2P diverges out from S_1 and S_2 and they superimpose thereby forming interference fringes in the overlapping region. In between region PQ, we can observe interference fringes with equal width and beyond that fringes are of unequal width due to the diffraction effects coming into picture.

3.6 Interference by Division of Wavefront: Young's Double Slit Experiment – Analytical Treatment

In early 1800, British scientist Thomas Young performed the first detailed experiment to analyze the variation of intensity on the screen due to the presence of two light waves on the screen. In fact, it was the first experimental evidence of wave theory of light, explaining the interference of light which could not be explained of the basis of ray optics.

Consider, S_1 and S_2 as two slits which fall on the same wavefront of a highly monochromatic and coherent source of light S as shown in Fig.3.7.

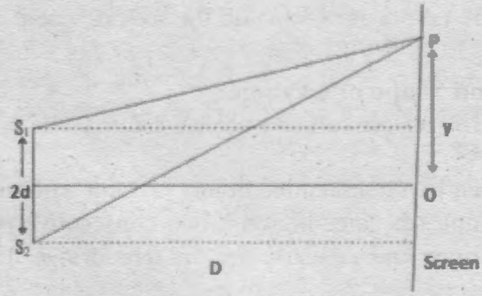


Fig. 3.7. Young's Double Slit Experiment

Let D is the distance of the screen from the two sources and 2d is distance between two sources/slits. P is an arbitrary point on the screen where the intensity of light is to be analyzed under the influence of two light waves reaching from two slits.

From equation (9), we have seen that the intensity variation is due to different phase difference at different points on the screen and phase difference is introduced due to this path difference. We have to thus find out the path difference between the two interfering waves, in order to get intensity at various points on the screen. According to the Fig.3.7, let the two waves S_1P and S_2P meet at an arbitrary point P on the screen which is at a distance y from the centre O, of the screen. From Fig. 3.7, the path difference can be calculated as;

$$
\begin{aligned}
S_2P - S_1P &= \sqrt{D^2 + (y+d)^2} - \sqrt{D^2 + (y-d)^2} \\
&= D\left[\left\{1 + \left(\frac{(y+d)^2}{D^2}\right\}^{1/2} - \left\{1 + \left(\frac{(y-d)^2}{D^2}\right\}^{1/2}\right.\right.\right] \\
&= \frac{2yd}{D} \qquad (21)
\end{aligned}
$$

3.6.1 Conditions for Maxima and Minima

(i) Maxima

$S_2P - S_1P = n\lambda$ for maximum intensity on the screen, i.e.

$\frac{2yd}{D} = 2n\,\lambda/2,$

If y_n is the distance of n^{th} maxima from the centre of the screen, then;

$y_n = \frac{nD\lambda}{2d}$ are the points of maxima, where n= 0,1,2... (22)

At the centre of the screen, the path difference / phase difference is zero which is the condition for maximum intensity. It means maxima will be formed at the centre of the screen but zero order minima has no physical significance.

(ii) Minima

For minimum intensity at the point P, called as destructive interference, the path difference must be odd multiple of $\lambda/2$, i.e.

$S_2P - S_1P = (2n \pm 1)\,\lambda/2$ for minimum intensity on the screen, i.e.

$\frac{2yd}{D} = (2n \pm 1)\,\lambda/2,$

If y_n is the distance of n^{th} minima from the centre of the screen, then;

$y_n = \frac{(2n\pm1)D\lambda}{4d}$ are the points of minima on the screen, where n = 1, 2, 3... (23)

3.6.2 Fringe, Fringe Width and Shape of a Fringe

Fringe: Regular bright and dark regions on the screen created due to constructive and destructive interference are called fringes.

Fringe Width: The gap between two consecutive points of same intensity is one fringe and this width is called fringe width. For example: distance between two consecutive points of maximum intensity is called fringe width of bright fringe and similarly for the dark fringe. Fringe width of bright or dark fringe comes out to be:

$\omega = y_{n+1} - y_n = \frac{D\lambda}{2d}$ for both bright and dark fringe.

This shows that fringe width is same and constant for bright and dark fringes.

Shape of a Fringe: The locus of constant path difference, traces the shape of a fringe.

3.6.3 Outcome of The Young's Double Slit Experiment

The physical phenomena obtained by meeting of two or more than two light sources could not be explained on the basis of ray theory of light. Young's double slit experiment is the first experiment to explain successfully the superposition of light rays as well as to prove the wave nature of light calling them two light waves instead of light rays. By proving the wave nature of light, various phenomenon like diffraction, polarization, scattering etc. along with the electromagnetic nature of light, could be explained leading to enormous applications of light.

3.7 Young's Double Slit Experiment Using Fresnel Bi-Prism

The two slits in the analytical treatment of YDSE should act as two coherent sources of light sending light waves. But we have already discussed that two independent sources cannot be coherent. Thus, for theoretical point of view, the analytical treatment of YDSE, is acceptable but cannot be performed practically. For practical purpose, we use division of wavefront method to produce two coherent sources using Fresnel biprism. Hence, from measurement point of view the Young's double slit experiment is done using Fresnel biprism.

This section comprises of a complete application of YDSE to determine the wavelength of monochromatic light using Fresnel biprism. The experimental set up is already shown in Fig.3.7 above. All the steps followed for the derivation of various formulae will be applied for Fresnel Biprism also, as used in the analytical treatment of YDSE (see section 3.6). Using the following formula for fringe width, we can determine the wavelength of monochromatic light.

$$\omega = y_{n+1} - y_n = \frac{D\lambda}{2d} \quad \text{or} \quad \lambda = \omega . \frac{2d}{D} \qquad (25)$$

Where ω is the width of a fringe,
D is the distance of two sources from the screen,
2d is the distance between the two sources,
λ is the wavelength of monochromatic light used.

With the help of formula (25), the wavelength of monochromatic light can be measured. In other words, in order to find the wavelength of light, we need to measure the distance of two sources from the screen i.e., D, fringe width ω and 2d i.e., the distance between two virtual sources.

Adjustment of the Experimental Set-Up

A narrow adjustable slit and the bi-prism are placed on the optical bench in such a way that the refracting edge of the bi-prism and the length of the slit are parallel and on the same height. Due to line source, a cylindrical wave front impinges on the two refracting faces of the two prisms. Each portion of the wave front is deviated towards the common base of the two prisms, appearing as if the two wave fronts are generated from the two sources. In fact, single wave front is divided into two parts that is why this is called generation of two sources by the division of wave front. By the proper alignment of slit, bi-prism and eyepiece on the optical bench, a sustainable and well-defined interference pattern can be obtained on the screen (Fig. 3.8). The various measurements can be taken up as:

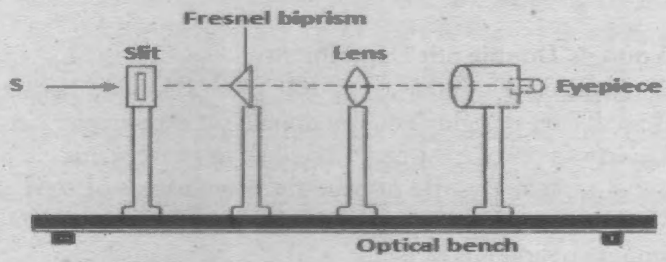


Fig.3.8 Experimental set up for determination of wavelength by Fresnel's biprism

1. Measurement of D: As shown in the diagram, D the distance of two sources from the screen can be measured from the optical bench.
2. Measurement of fringe width: The cross wire of the eyepiece is to be set on any particular fringe, in the pattern obtained on the screen and the reading is taken then on the scale fixed with the eyepiece. The eyepiece is then moved on the either side of the pattern and the fringes are then counted that are crossed in moving the cross wire on any same kind of fringe (bright/ dark). Again, the readings are taken on the scale. Taking the difference between the two readings and by dividing by the number of fringes which are crossed in between the two, we get the width of a fringe.
3. Measurement of 2d: There are two methods to find the distance between the two virtual sources. One is purely based on the measurements and the other one, in which we have to rely on the supplier for the physical parameters of the bi-prism.

(a) First Method involves determination of 2d with the help of a convex lens. A convex lens is placed between the slit and eyepiece and the image of two sources is seen through the eyepiece Fig.3.9. Now adjusting the cross wire of the eyepiece on the images of the two sources, we find the separation between them. Take this separation as the height of the image, let it be d_1. Now keeping the position of slit, bi-prism and screen unchanged, move the position of the lens in such a way that we again get sharp images of the two virtual sources through the eyepiece.

Again, measure the separation between the two images with the help of cross wire of the eyepiece, let this be d_2. Using the magnification formula:

$$\frac{v}{u} = \frac{I}{O},$$

If 2d is considered as the height of object, then

$$\frac{v}{u} = \frac{d_1}{2d}, \quad \text{for first position of the lens,} \tag{26}$$

$$\frac{v}{u} = \frac{d_2}{2d}, \quad \text{for second position of the lens,} \tag{27}$$

From (26) and (27) we get;

$$2d = \sqrt{d_1 d_2} \tag{28}$$

Now using the formula (25), wavelength can be measured. In this method, the error in the measurement of wavelength may arise due to the errors in the measurement involved for different physical parameters. These can be minimized by minimizing the errors in the various measurements.

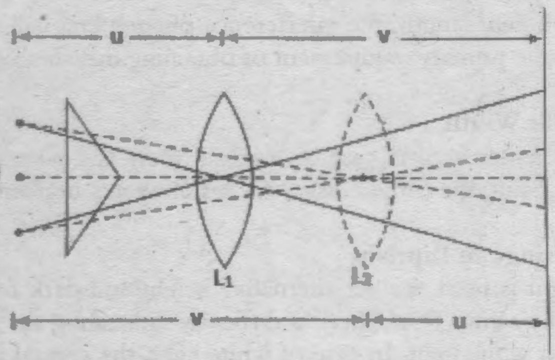


Fig. 3.9 Measurement of distance 2d between virtual sources using two different positions of convex lens

(b)Alternative Method: If the angle of prism is very small, then in order to minimize the deviation and ultimately to keep the distance between two sources small,

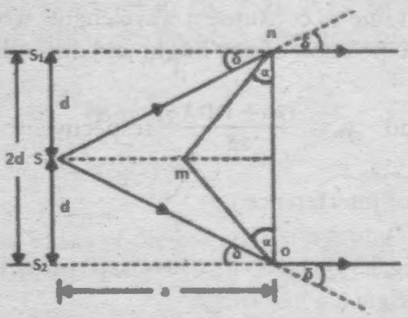


Fig. 3.10. Measurement of distance 2d between virtual sources using deviation method

$2d = 2a\,(\mu - 1)\,\alpha$ (see Fig. (3.10))

Where; a = the distance between sources and bi-prism (measured)

μ = refractive index of the material of the bi-prism (given)

α = is the angle of the bi-prism (given)

By putting the measured value of a and given values of μ, α, we can find 2d and then finally the wavelength of the light used. In this method, most of the values of the physical parameters are given. If these values are not correct for a particular wavelength, there may be errors in the measurement in finding the wavelength and hence cannot be minimized. This suggests that from experimental point of view, first method is advisable.

Need For Small Angle of Bi-Prism

As the angle of prism is directly proportional to the deviation produced by the prism. The large deviation produced by the biprism will lead to a very small region available for interference. This is one of the drawbacks for large angle of bi-prism. Another drawback can be seen directly from the equation $2d = 2a\,(\mu-1)\,\alpha$. According to this, large α leads to large 2d, i.e., large distance between two sources. The effect of large distance between two sources can be understood in many ways;

(i) Large 2d leads to very small fringe width resulting in non- differentiable fringes.

(ii) Large 2d leads to large path difference between the two interfering ways. If the path difference exceeds the temporal coherent length, the interference phenomena will not be consistent. This will then be contradictory to the primary requirement of obtaining measurable interference pattern.

Effect of Increasing the Slit Width

When the width of the slit is increased in case of biprism, there is a poor contrast of bright and dark interference fringes. At one point, the fringes disappear resulting in a uniform illumination everywhere.

Location of Zero Order Fringe in Biprism

When a monochromatic light is used, we get alternative bright and dark fringes all of equal thickness which physically appear to be same. Thus, there is difficulty in locating the zero-order fringe. This can be overcome with the use of white light. In case of white light, the central fringe is white and all other fringes coloured. The vertical crosswire is then adjusted on the zero order or the central white fringe and white light is again replaced by a monochromatic light. Thus, zero order can be located.

3.8 Fringes with White Light Using a Biprism

White light consists of wavelengths ranging from 4000 - 7000 Å. When white light (e.g., sunlight) is made to fall on the slit, the effect due to constituent wavelengths would be different as compared to the monochromatic light if used. The position of any bright and dark fringe can be given as:

$$y_n = \frac{nD\lambda}{2d} \quad \text{and} \quad y_n = \frac{(2n \pm 1)D\lambda}{2d} \quad \text{respectively.}$$

n is an integer, known as order of interference.

(i) The point O (as shown in Fig.3.11) is a point where zero order bright fringe lies i.e. where $y = 0$ for all the wavelengths. Thus, we get a white fringe.

(ii) Fringe width is given as, $w = \frac{D\lambda}{2d}$ which is directly proportional to λ. Red with greatest wavelength will have greater fringe width in comparison to violet with lesser wavelength and hence lesser fringe width. The fringe width for violet rays is approximately half the fringe width for red light.

(iii) The path difference goes on increasing as we move on either side of O. At any point X away from the centre O, the path difference, $\Delta = S_2P - S_1P$ is same for all the wavelengths but contrary to this phase difference, $\delta = \frac{2\pi}{\lambda}\Delta$ is not the same for all the wavelengths. As a result, some of the points may have maximum intensity for a number of wavelengths for which $\Delta = n\lambda$ and minimum intensity for which $\Delta = (2n-1)\lambda/2$. There will be then overlapping of intensity due to various wavelengths and point X will appear to be coloured as long as Δ is small.

(iv) The path difference becomes larger when point X is farther away from the centre. Thus, a large number of wavelengths will produce maximum intensity and equally large number produce minimum intensity at X. This will lead to the general illumination.

We can conclude as: When white light illuminates the slit, an interference pattern is observed with central bright fringe with few coloured fringes symmetrically situated on both the sides followed by general illumination.

3.9 Shifting of Fringe Pattern

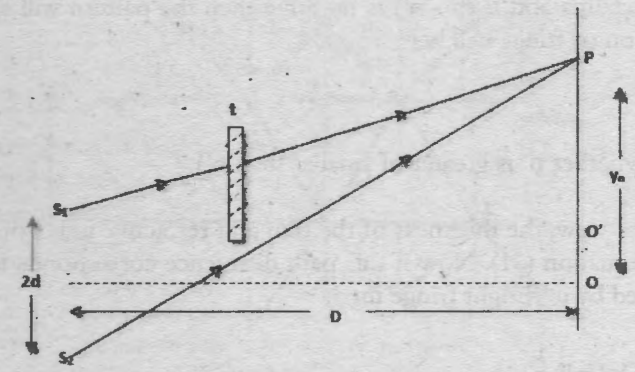


Fig. 3.11 Displacement of fringes using transparent sheet

In Fresnel Bi-Prism experiment, it can be seen that the position of a particular fringe is shifted towards or away from the central fringe by introducing a transparent sheet (e.g., glass, mica etc.) in the path of one of the interfering beams of light as shown in Fig. 3.11. It is due to the fact that the position of a particular fringe is related to the path difference between the interfering rays. The position of any bright or dark fringe can be given as:

$y_n = \frac{nD\lambda}{2d}$, (Bright),

$y_n = \frac{(2n\pm1)D\lambda}{4d}$ (Dark)

If a film of thickness t and refractive index of the material of the film μ is introduced in the path of one of the interfering waves, then the path difference between the two waves at the same point will be;

$$\Delta = S_2P - [S_1P + (\mu-1)t]$$
$$= S_2P - S_1P - (\mu-1)t \qquad (29)$$

Now if this path difference corresponds to any bright fringe, then

$$S_2P - S_1P - (\mu-1)t = n'\lambda$$

Or $\quad n\lambda - (\mu-1)t = n'\lambda$

Where $n\lambda = S_2P - S_1P$ is the path difference corresponding to n^{th} bright fringe. So, the number of fringes shifted will be given as:

$$(n - n') = \frac{(\mu-1)t}{\lambda}, \qquad (30)$$

and the displacement on the screen will be

$$(n - n')\,\omega = \frac{\omega\,(\mu-1)t}{\lambda} = \frac{D\,(\mu-1)t}{2d} \qquad (31)$$

Where, ω is the fringe width. From (31) it can be seen that if (n - n′) is positive, the pattern will shift far away from the central maxima and if (n - n′) is negative then the pattern will shift towards the central maxima. And new position of fringe will be:

$$y_{n'} = \frac{nD\lambda}{2d} \pm \frac{D(\mu-1)t}{2d} \qquad (32)$$

(+ or - depends upon whether n is greater of smaller than n′)

From application point of view, the thickness of the film and refractive index of the material of the film can be calculated using equation (31). Now if this path difference corresponds to any dark fringe at the point, previously occupied by n^{th} bright fringe then;

$$S_2P - S_1P - (\mu - 1)t = (2n'-1)\frac{\lambda}{2}$$

Or $$n\lambda - (\mu - 1)t = (2n'-1)\frac{\lambda}{2}$$

$$(n - n')\lambda = (\mu - 1)t - \frac{\lambda}{2}$$

Or $$(n - n') = \frac{(\mu - 1)t}{\lambda} - \frac{1}{2}$$

And the displacement of the pattern will be given as;

$$(n - n')\omega = \frac{\omega(\mu - 1)t}{\lambda} - \omega/2 \quad \text{(where } \omega \text{ is the fringe width)} \qquad (33)$$

3.10 Thin Film Interference

To observe the phenomenon of interference by the division of wavefront, it is necessary to use a narrow source as we have seen already while studying YDSE and Fresnel biprism. On the contrary, in order to study thin films e.g.be it be a thin film of an oil spread on the surface of water or a thin film of transparent material like soap bubbles etc. based on the principle of division of amplitude, we need to use an extended or a broad source. As can be seen from the Fig. 3.12, the interference in thin films arises due to superposition of waves arising from multiple reflection and refraction of light from the upper and the lower surfaces of the film.

3.10.1 Interference in Reflected Light

Consider a thin parallel transparent film with thickness t, refractive index $\mu>1$. Let a monochromatic light SA of wavelength λ falls from a medium such as air ($\mu = 1$) on the upper surface of the thin film at an angle of incidence i. This ray gets partly reflected along AP and partly refracted along AB at an angle r. At B again it is partly reflected from the lower surface of the film along BC and emerges out as CQ as shown in Fig. 3.12. The rays AP and CQ both undergo one reflection, have practically the same intensity; although the intensity of directly and internally reflected rays goes on decreasing rapidly. We can see that both of these are actually derived from the same incident beam SA and hence are said to be coherent. The rays AP and CQ are in a position to interfere as the two rays are sufficiently close to each other due to a thin film. Since the two surfaces of the film are parallel, the rays AP and CQ are parallel to each other.

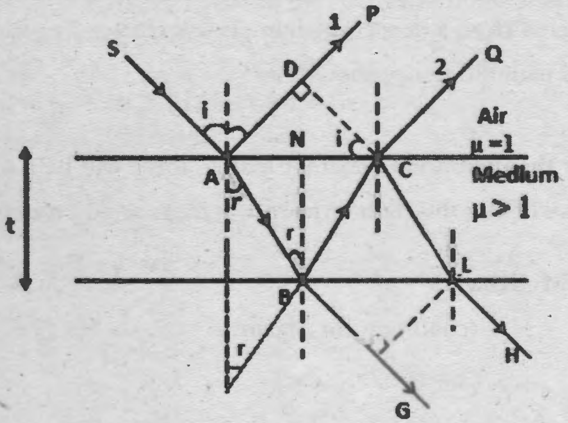


Fig.3.12. Interference in reflected region using thin parallel film

The interference of these two rays will be constructive or destructive, depending on the phase difference between them. Geometrically, path difference measurement is the easiest way to find the phase difference between the two wave trains. In this case, the light has to travel in two media, so it is better to find optical path difference between them. As μ, r and t are the characteristic parameter of a particular film, so we have to find the path difference in terms of these parameters only. Before we calculate it draw CD normal to AP and BN normal to AC. As the paths of the rays AP and CQ beyond CD are equal, the optical path difference is given as:

Optical path difference, d = path ABC in film – path AD in air

$$= \mu\,(AB + BC) - AD \qquad (34)$$

In right triangle ABN, $\cos r = \frac{BN}{AC}$

$$AC = \frac{t}{\cos r} \qquad (BN = t) \qquad (35)$$

Also, in right triangle BCN, $\cos r = \frac{BN}{BC}$

$$BC = \frac{t}{\cos r} \qquad (36)$$

So, In rt.Δ ACD, $\sin i = \frac{AD}{AC}$

Thus, $$AD = AC \sin i = (AN + NC)\sin i \qquad (37)$$

But in Δ ANB, AN = BN tan r = t tan r

In Δ BNC, NC = BN tan r = t tan r

From equation (37),

$AD = (t\tan r + t\tan r)\sin i = 2t\tan r\sin i$

$$= 2t\frac{\sin r}{\cos r}\sin i = 2t\frac{\sin r}{\cos r}\mu\sin r = 2\mu t\frac{\sin^2 r}{\cos r} \quad \{\sin i = \mu\sin r,\ \text{Snell's law}\} \qquad (38)$$

Substitute the values of AC, BC and AD from (35), (36) and (38) in equation (34), we get

$$d = \left[\frac{t}{\cos r} + \frac{t}{\cos r}\right] - 2\mu t\frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r} - 2\mu t\frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r}\left[1 - \sin^2 r\right]$$

$$= \frac{2\mu t}{\cos r}\cos^2 r = 2\mu t\cos r \qquad (39)$$

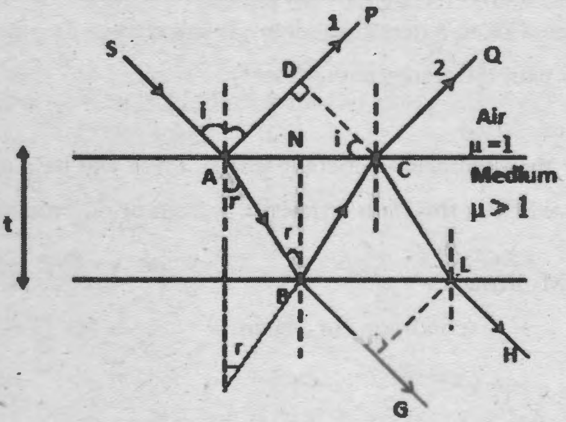


Fig.3.12. Interference in reflected region using thin parallel film

The interference of these two rays will be constructive or destructive, depending on the phase difference between them. Geometrically, path difference measurement is the easiest way to find the phase difference between the two wave trains. In this case, the light has to travel in two media, so it is better to find optical path difference between them. As μ, r and t are the characteristic parameter of a particular film, so we have to find the path difference in terms of these parameters only. Before we calculate it draw CD normal to AP and BN normal to AC. As the paths of the rays AP and CQ beyond CD are equal, the optical path difference is given as:

Optical path difference, d = path ABC in film – path AD in air

$$= \mu\,(AB + BC) - AD \qquad (34)$$

In right triangle ABN, $\cos r = \frac{BN}{AC}$

$$AC = \frac{t}{\cos r} \qquad (BN = t) \qquad (35)$$

Also, in right triangle BCN, $\cos r = \frac{BN}{BC}$

$$BC = \frac{t}{\cos r} \qquad (36)$$

So, In rt.Δ ACD, $\sin i = \frac{AD}{AC}$

Thus, $$AD = AC \sin i = (AN + NC)\sin i \qquad (37)$$

But in Δ ANB, AN = BN tan r = t tan r

In Δ BNC, NC = BN tan r = t tan r

From equation (37),

AD = (t tan r + t tan r)sin i = 2t tan r sin i

$$= 2t\frac{\sin r}{\cos r}\sin i = 2t\frac{\sin r}{\cos r}\mu \sin r = 2\mu t\frac{\sin^2 r}{\cos r} \qquad \{\sin i = \mu \sin r, \text{ Snell's law}\} \qquad (38)$$

Substitute the values of AC, BC and AD from (35), (36) and (38) in equation (34), we get

$$d = [\frac{t}{\cos r} + \frac{t}{\cos r}] - 2\mu t\frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r} - 2\mu t\frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r}[1 - \sin^2 r]$$

$$= \frac{2\mu t}{\cos r}\cos^2 r = 2\mu t \cos r \qquad (39)$$

But, the wave train 1 is reflected from a denser medium, it will change its phase by π. It will carry extra path equal to $\frac{\lambda}{2}$, so the actual path difference should be;

$$= 2\mu t \cos r - \frac{\lambda}{2}$$

If this path difference is less than temporal coherent length, there will be a maximum or minimum of interference depending upon whether this path difference is even or odd multiple of $\frac{\lambda}{2}$

Condition For Maxima or Minima

If $2\mu t \cos r - \frac{\lambda}{2} = 2n\frac{\lambda}{2}$ (condition for Maxima)

Or $2\mu t \cos r = (2n+1)\frac{\lambda}{2}$ (n= 0,1,2 .....etc.)

If $2\mu t \cos r - \frac{\lambda}{2} = (2n-1)\frac{\lambda}{2}$ (condition for Minima)

Or $2\mu t \cos r = 2n\frac{\lambda}{2}$ (n = 1,2,3....etc.)

From the above conditions, we can see that if the film is of uniform thickness and the incident wave front on the surface of the film is parallel, the whole film can either satisfy the condition for maximum or minimum for a particular wavelength. It means in monochromatic light; the film will look either dark or bright of that wavelength.

3.10.2 Interference in Transmitted Region

Consider a thin parallel transparent film with thickness t, refractive index $\mu > 1$. Let a monochromatic wave train SA of wavelength λ falls from a medium such as air ($\mu = 1$) on the upper surface of the thin film at an angle of incidence i. This ray after refraction follows path AB. Also, r as the angle of refraction for any angle of incidence i. At B, it is partly reflected along BC and partly reflected along BP. At C, the ray gets reflected along CF and part is refracted along FQ. As the rays 3 and 4 are both derived from the same incident wave, they are coherent see Fig.3.13. The rays BP and FQ are in a position to interfere as the two rays are sufficiently close to each other due to a thin film. Since, the two surfaces of the film are parallel, the rays BP and FQ are parallel to each other. Since both the reflections: at B and C takes place at the rarer medium, no phase change takes place.

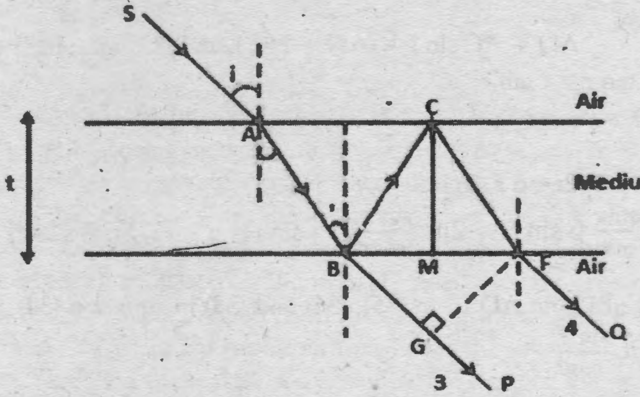


Fig.3.13. Interference in transmitted region using parallel thin film

If the necessary and sufficient conditions for the interference to take place, are satisfied in the transmitted region also, we can find conditions for maxima and minima in that region also. But let us now calculate the optical path difference between the two wave trains 3 and 4. Draw CM normal to BF and FG normal to BP. As the path of the rays BP and FQ beyond FG are equal, the path difference between the rays is given as:

Optical path difference, d = path BCF in film – path BG in air

$$= \mu (BC + CF) - BG. \qquad (40)$$

In right triangle BCM, $\cos r = \frac{MC}{BC}$

$$BC = \frac{t}{\cos r} \qquad (MC = t) \qquad (41)$$

Also, in right triangle MCF, $\cos r = \frac{MC}{CF}$

$$CD = \frac{t}{\cos r} \qquad (42)$$

In rt.Δ BFN, $\sin i = \frac{BG}{BF}$

Thus, $$BG = BF \sin i = (BM + MF)\sin i \qquad (43)$$

But in Δ BCM, BM = CM tan r = t tan r

In Δ CFM, MF = CM tan r = t tan r

From equation (43),

BG = (t tan r + t tan r) sin i = 2t tan r sin i

$$= 2t \frac{\sin r}{\cos r} \sin i = 2t \frac{\sin r}{\cos r} \mu \sin r = 2\mu t \frac{\sin^2 r}{\cos r} \quad \{\sin i = \mu \sin r, \text{ Snell's law}\} \qquad (44)$$

Substitute the values of BC, CF and BG from (41), (42) and (44) in equation (40), we get

$$d = [\frac{t}{\cos r} + \frac{t}{\cos r}] - 2\mu t \frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r} - 2\mu t \frac{\sin^2 r}{\cos r}$$

$$= \frac{2\mu t}{\cos r} [1 - \sin^2 r]$$

$$= \frac{2\mu t}{\cos r} \cos^2 r = 2\mu t \cos r \qquad (45)$$

Thus, we can examine clearly that, with the similar treatment as in case of reflected region, the geometrical path difference comes out to be $2 \mu t \cos r$. Since, both the wave trains 3 and 4 suffer either transmission or reflection from denser to rare interface and in both the cases there is no phase change. So, the actual path difference will be as it is. The conditions for maximum or minimum will be:

$2 \mu t \cos r = 2n\frac{\lambda}{2}$ (condition for Maxima)

Where n= 0,1,2............ etc.

$2 \mu t \cos r = (2n+1)\frac{\lambda}{2}$ (condition for Minima).

Where n = 0,1,2........... etc.

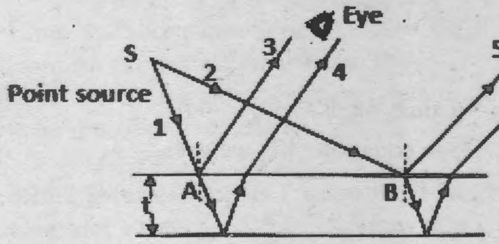
From the above conditions for maxima and minima in reflected and transmitted region, it can be seen that the conditions are complimentary i.e., the condition for maxima in the reflected region and minima in transmitted region is same and the condition for minima in reflected and maxima in transmitted region is same. In other words, we can say that if the film is illuminated with white light, the colour which is present in reflection, it will be absent in transmission and vice a versa. Here, we see that a film of a particular set of physical parameters (μ, r and t) can filter a particular colour from a band of colours (e.g,, white light) of the film.

Colours in Reflected and Transmitted Light are Complementary

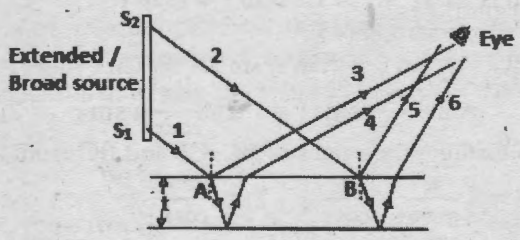
As we have already shown with the help of derivation that the conditions of maxima and minima for reflected light are opposite to those for transmitted light; thus, with the use of white light, the colours visible in the reflected light will be complementary to those visible in transmitted light. This is because if the thickness of the film is such that the condition of maxima holds for certain colours in reflected system, the conditions of minima will hold for the same colours in the transmitted system and vice a versa. This implies the colour absent in one system will be present in the other.

Why to choose a broad source for viewing thin films?

We know already that a narrow source is required to obtain interference fringes in the case of Fresnel's biprism, Lloyd's single mirror etc. by two coherent sources. The fringes thus obtained on the screen can be viewed with the help of an eyepiece. But the use of narrow source limits the visibility of the film in the case of interference in thin films. An extended source is therefore necessary to enable the eye to see whole of the film simultaneously. This may be understood as follows:



(a) Narrow source



(b) Extended source

Fig. 3.14.

Consider a narrow source S through which light is incident on a thin film. Also, 1,2 represents the two incident rays. For each of these incident waves we get a pair of parallel interfering waves. The ray 1 produces interference fringes because 3 and 4 reach the eye whereas the ray 2 meets the surface at some different angle and is reflected along 5,6 and thus do not reach the eye as shown in Fig.3.14 (a)Similarly, the other rays can be incident on the surface at different angles which may not reach the eye. Due to the limited size of pupil, the rays from only a small area of the film can enter the eye. Thus, only a small portion of the film will be visible. To observe the different parts of the film, the eye has to be moved throughout. Hence, with a narrow source it is impossible to observe the entire film altogether. If an extended source of light is used, the ray 1 after reflection from the upper and the lower surface of the film emerges as 3 and 4 which reach the eye. Also ray 2 from some other point of the source after reflection from the upper and the lower surfaces of the film emerges as 5 and 6 which also reach the eye as shown in Fig. 3.14 (b). Therefore, in the case of an extended source of light, the rays incident at different angles on the film are accommodated by the eye and this enables the eye to view a larger area of the film. Due to this reason, to observe interference phenomenon in thin films, an extended or a broad source of light is required. With a broad source of light, rays of light are incident at different angles and the reflected parallel beams reach the eye placed in a suitable position. Each such ray of light has its origin at a different point on the source. Although any extended source can be employed here; but the conditions of interference are not violated.

3.11 Colours in Very Thick, Thin Films and Different Shaped Films

(i) Very Thick Film: If white light is used to illuminate a film having thickness exceeding even a few wavelengths of light, no interference will be observed. In this case the condition of constructive interference, i.e., $2\mu t \cos r = (2n - 1)\lambda/2$, at a given point is satisfied by a large number of wavelengths (colours) [the value of n is different for different colours] and at the same point the condition of destructive interference, i.e., $2\mu t \cos r = n\lambda$, is satisfied for another set of large number of wavelengths (colours). Further the number of wavelengths sending maximum intensity at a given point is almost equal to the number of wavelengths sending minimum intensity. Moreover, the fringes due to wavelengths satisfying the condition of maximum will overlap and produce a uniform white illumination. Thus, in the case of a thick film illuminated by white light, the colours are not observed in the reflected light.

Note: The large thickness of the film leads to very large path difference between the two interfering waves that may extend to the temporal coherent length. Due to this temporal incoherency, the interference between the waves from front and back surface will not be consistent. So, whether the light is coloured or monochromatic, the interference will not be seen in case of a very thick film.

(ii) Very Thin Film: If we have an excessively thin film such that $2\mu t \cos r \ll \lambda/2$ it appears black in reflected light. We know that the effective path difference in the reflected light for a film of refractive index μ and thickness t between the rays reflected from the upper and lower surfaces of the film is given by:

$$2\mu t \cos r + \lambda/2$$

where r is the angle of refraction in the film and λ is the wavelength of light used. For an excessively thin film whose thickness is small in comparison with the wavelength of light used, then from the above expression $2\mu t \cos r$ will be negligible in comparison to $\lambda/2$. Thus, the effective path difference becomes $\lambda/2$. This is a condition for minimum intensity. This is true for all wavelengths. Each of the wavelength will be absent in the reflected light and hence the film will appear black in reflected light even when illuminated with white light.

(iii) In Different Thin films (wedge or parallel shaped soap bubble/thin layer of oil film): A thin film or a soap bubble or a thin layer of an oil film is illuminated with white light and can be viewed at different angles. As the different angles of inclination alter the path difference and the different path difference may correspond to the condition for maximum for different wavelengths, the film could be seen as coloured.

3.12 Interference with Wedge Shaped Thin Film

If the thickness of the film is not uniform and varies from one side to other keeping the two-surface plane, this shape is termed as wedge shaped thin film. If the angle of wedge is very small, the interference phenomena will be observed in the region of the film where the path difference between the reflected wave trains in the reflected region or transmitted wave trains in the transmitted region is less than the temporal coherent length of the coherent beam of light. Let a wedge-shaped thin film with refractive index of the material as μ and angle of wedge θ, is placed in air. The two plane surfaces OP and OQ are inclined at an angle θ. A beam of light falls on it at an angle of incidence for which r is the angle of refraction, and thickness of the film at that point is t. The thickness increases from O to P shown in Fig.3.15. The film when illuminated by a monochromatic source of light from an extended source, gives straight equidistant alternate bright and dark fringes which are parallel to the edge of the wedge passing though O. This is because of the interference between the rays reflected from upper and the lower surfaces of the film. The rays are coherent as they are originated from the same source. Let these rays be AB and DE; both originating from the same incident ray SA. Draw the perpendiculars DK and DN from D on AB and AC respectively. Now let us calculate the path difference between the required two wave trains in reflected and transmitted region.

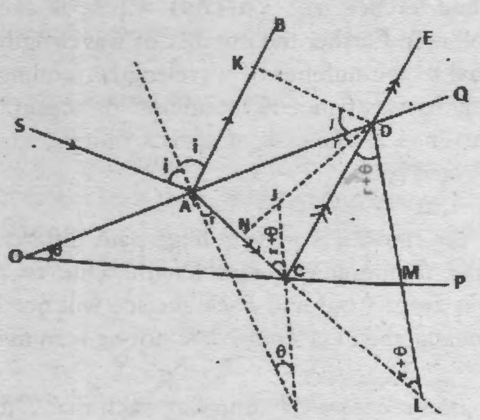


Fig. 3.15. Wedge shaped film

Now, the path difference in reflected region will be calculated as follows:
As we know that the path difference should be in terms of physical parameters of the particular film (μ, r, t and θ), so let us relate AB, AC and AD with these parameters.

$$p = (AC + CD)\text{in film} - AK \text{ in air}$$
$$= \mu(AC + CD) - AK \qquad (46)$$
$$= \mu(AN + NC + CD) - AK$$

From simple geometry it can be shown that

$$\angle ADK = i \text{ and } \angle ADN = r$$

$$\therefore \qquad AK = \mu.AN$$

Substituting this value of AK in equation (1), we get

$$p = \mu(AN + NC + CD) - \mu.AN$$
$$= \mu.NC + \mu.CD \qquad (47)$$

Now draw perpendicular DM from D on OP and produce AC. These meet at L.

Now triangles CDM and CML are congruent,

Since $\angle CDM = \angle CLM = r + \theta$

$\angle DMC = \angle CML = 90°$

and CM is common

$$\therefore \qquad DM = Ml = t$$

and $\qquad CD = Cl$

Substituting value of CD in equation (47), we get

$$p = \mu.NC + \mu.CL = \mu(NC + CL) = \mu.NL \qquad (48)$$

As the angle enclosed between the surfaces OP and OQ is θ, the angle enclosed between normals to the surfaces at A and C must be θ. The angle of incidence ACJ at C is therefore $(r + \theta)$. CJ and DL are normals to the surface OP; therefore, CJ and DL are parallel. As ACL cuts the parallel lines CJ and DL, we must have

$$\angle ACJ = \angle CLD = (r + \theta).$$

In right angled triangles NDL,

$$NL = DL\cos(r + \theta) = 2t\cos(r + \theta).$$

Equation (48) gives,

$$p = 2\mu t\cos(r + \theta).$$

As the wave train along AB is the reflected wave train from a denser medium; therefore, there occurs a phase change of π or path difference of $\frac{\lambda}{2}$.

The effective path difference between the interfering waves AB and DE is given by

$$\Delta = 2\mu t\cos(r + \theta) - \frac{\lambda}{2} \tag{49}$$

Conditions For Interference

For constructive interference

$$2\mu t\cos(r + \theta) - \frac{\lambda}{2} = 2m\frac{\lambda}{2}$$

or $$2\mu t\cos(r + \theta) = (2m + 1)\frac{\lambda}{2}, \qquad m = 0,1,2,3 \ldots.$$

or $$2\mu t\cos(r + \theta) = (2n - 1)\frac{\lambda}{2}, \qquad n = 0,1,2,3 \tag{50}$$

For destructive interference

$$2\mu t\cos(r + \theta) - \frac{\lambda}{2} = (2n - 1)\frac{\lambda}{2}$$

or $$2\mu t\cos(r + \theta) = n\lambda \ldots\ldots\ldots\ldots\ldots n = 0,1,2,3 \ldots \tag{51}$$

Nature of Interference Pattern

When the wedge-shaped film is illuminated by a parallel beam of monochromatic light, the angle of incidence and hence the angle of refraction is constant at every point of the film. In this case μ, r and θ are constants and hence the path difference $2\mu t\cos(r + \theta)$ will change due to varying thickness of the film from point to point. But for a bright or dark fringe of any particular order the path difference must be constant. For this t must remain constant. Since the locus of points of constant thickness of the film is a straight line parallel to the edge of the film, the fringes are straight lines parallel to the edge.

The film will appear bright if the thickness t of the film satisfies the condition (50), i.e.

$$2\mu t\cos(r + \theta) = (2n - 1)\frac{\lambda}{2}$$

$$t = \frac{(2n-1)}{4\mu\cos(r + \theta)}, \qquad n = 0,1,2,3 \ldots.$$

$$t = \frac{\lambda}{4\mu\cos(r + \theta)}, \frac{3\lambda}{4\mu\cos(r + \theta)}, \frac{5\lambda}{4\mu\cos(r + \theta)} \ldots \text{ etc.}$$

On the other hand, the film will appear dark if satisfies condition (51) i.e., $2\mu t\cos(r + \theta) = n\lambda$

$$t = \frac{n\lambda}{2\mu\cos(r + \theta)}, \qquad n = 0,1,2,3 \ldots.$$

$$t = 0, \frac{\lambda}{2\mu\cos(r + \theta)}, \frac{2\lambda}{2\mu\cos(r + \theta)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Thus, we find that if we move along the film in the direction of increasing thickness, we shall observe alternatively dark and bright fringes parallel of the edge of the film.

Why the edge of the film appears dark in case of wedge shape films?

At the edge, $t = 0$, therefore the path difference is $\lambda/2$, a condition for darkness. Therefore, the edge of the film appears dark. This is called zero order band. Beyond the edge, for thickness

$$t = \frac{\lambda}{4\mu\cos(r+\theta)}$$

The path difference is λ and we obtain first bright band. As t increases to a value $\frac{\lambda}{2\mu\cos(r+\theta)}$, the path difference is $\frac{3\lambda}{2}$ and we obtain first dark band. Thus, as thickness increases, we obtain alternate dark and bright bands.

3.13 Shape of Fringe Pattern in the Interference Phenomena

(i) Shape of a Fringe in YDSE: Shape of any fringe is the locus of constant path difference i.e., locus of a particular intensity pattern (bright or dark) at different location of the screen. In case of Young's Double Slit experiment, the path difference for n^{th} bright fringe is given by;

$S_2P - S_1P = n_1 \lambda$

And for n^{th} dark fringe, the path difference is

$S_2P - S_1P = (2n \pm 1)\lambda/2$

In both the cases, for a particular fringe i.e., for constant n, the value of $S_2P - S_1P$ will be constant for different locations of the screen from the sources. But the position of a particular fringe on the screen will be changing and if we trace the position of a particular fringe with the change in the position of the screen then its curve is traced as shown in the Fig.3.16. Obviously, it is a hyperbola. But this hyperbola cannot be seen on the screen due to obvious reason because the plane of the screen and the hyperbola are perpendicular to each other. And we will see the intersection of these two, which looks like a straight line.

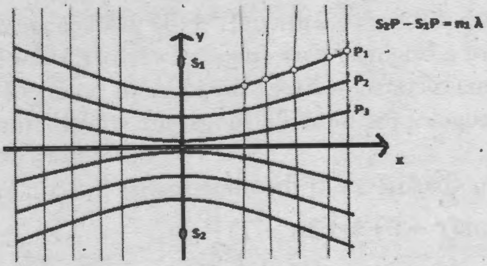


Fig. 3.16. Hyperbolic fringes

(ii) Shape in Parallel Thin Film: We know that the shape of a fringe is defined by the locus of position of fringes of constant path difference. In the interference phenomena occurring through a parallel thin film, the locus of constant path difference is a circle in the reflected and transmitted region.

(iii) Shape in Wedge Shaped Thin Film: constant path difference will trace a line parallel to the wedge, so fringe shape will be straight parallel to the edge of the wedge (Fig.3.17).

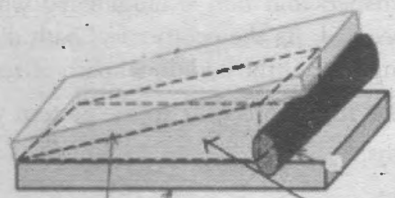


Fig. 3.17 Fringe Shape in Wedge Shaped Thin Film

3.14 Fringe Width in Wedge Shaped Thin Film: As the fringes are straight lines parallel to the edge of the wedge, we can find the spacing between two consecutive bright or dark fringes (known as fringe width) shown in Fig. 3.18

The condition for bright fringe of n^{th} order is given as;

$$2\,\mu t_n \cos(r+\theta) = (2n+1)\,\frac{\lambda}{2}$$

Similarly, for $(n+1)^{th}$ order, we have

$$2\,\mu\, t_{n+1} \cos(r+\theta) = (2n+3)\,\frac{\lambda}{2}$$

From the diagram, we have,

Fig.3.18.

$$\tan\theta = \frac{t_n}{x_n} = \frac{t_{n+1}}{x_{n+1}}$$

Putting the values of t_n and t_{n+1} in the equations, we have

$$2\,\mu x_n \tan\theta \cos(r+\theta) = (2n+1)\,\frac{\lambda}{2}$$

$$2\,\mu x_{n+1} \tan\theta \cos(r+\theta) = (2n+3)\,\frac{\lambda}{2}$$

On Subtracting the above two equations we get;

$$2\,\mu(x_{n+1} - x_n) \tan\theta \cos(r+\theta) = \lambda$$

Here, $(x_{n+1} - x_n)$ gives the distance between two consecutive bright fringes i.e., fringe width. So, the;

Fringe width, $\omega = (x_{n+1} - x_n) = \dfrac{\lambda}{2\,\mu \tan\theta \cos(r+\theta)}$

This expression is independent of n, so it may be said that the distance between two consecutive bright fringes is constant i.e. fringe width is constant. Similarly, it can be calculated for dark fringes also which will be the same as for bright fringes.

For normal incidence and small angle of wedge,

$\cos(r+\theta = 1)$ and $\tan\theta = \theta$, the fringe width can be calculated as;

$$\omega = \frac{\lambda}{2\,\mu\theta} \qquad (52)$$

Wedge Shaped Thin Film in White Light: If a wedge-shaped thin film is illuminated with white light, coloured straight fringes parallel to the edge will be observed. As the geometrical path difference between two interfering waves will be zero for all the wavelengths, but the additional path difference of $\frac{\lambda}{2}$, for every colour will give destructive interference at the edge for each colour so the pattern will start with black fringe followed by colours of increasing wavelengths.

3.15 Newton's Rings

This was first studied by Newton and hence the name Newton Rings. It is an example of interference in an air film with varying thickness. When a plano convex lens of large focal length is placed in contact with a plane glass plate, we get a thin film between the lower surface of the lens and the upper surface of the plate. At the point of contact, thickness is very small. With monochromatic light falling on such kind of films normally, we get a dark spot at the centre with alternate bright and dark circular rings when seen in reflected light and vice a versa in the transmitted light. If on the other hand, if we use white light, a series of concentric coloured rings are observed at the centre in both reflected as well as in transmitted. When the air film is wedge shaped and loci of points of equal thickness are circles concentric with the point of contact, the interference fringes or rings obtained are of equal thickness known as Newton's rings.

3.15.1 Newton's Rings Experiment

This experiment is one of the most important applications of the interference through thin films. With this experiment, we can find the wavelength of monochromatic light, calculate and compare the refractive indices of liquid/liquids etc. The mathematical treatment in this experiment follows the theory of interference through wedge shaped thin films. By doing some modifications and approximations, the appropriate formulations are obtained as follows;

Experimental Arrangement

As seen from the Fig.3.19., a plano-convex lens L of large radius of curvature is placed on an optically thick glass plate G. Due to large radius of curvature, the free space between convex side of the lens and the glass plate serves as air filled wedge shaped thin film. The thickness of the air film increases from point of contact of lens to outer edge. We can consider a concave lens shaped air film between convex lens and glass plate. If the combination is placed in any liquid, a concave lens shaped air film of the liquid will be obtained. The thickness of the film at the point of contact can be treated as zero. Now the monochromatic light from an extended source is made to fall perpendicularly on the combination of lens and glass plate, the reflected wave trains from the upper and the lower surface of the air film interfere in the reflected region and transmitted wave trains interfere in the transmitted region. As the loci of the constant path differences are the concentric circles, the fringes obtained are in the shapes of the concentric rings with center as the point of contact of the lens and plate. These are called Newton's Rings. Now we can establish mathematical relations between the diameter of a circular ring (fringe) and the wavelength of light, refractive indices of the concave lens shaped film between plano-convex lens and glass plate. These relations can be utilized to find out wavelength of light and the refractive index of any optically transparent liquid.

In the given experimental arrangement, S is an extended source of light placed at the focus of a convex lens. Parallel rays of light from the lens falls on the glass plate B at 45°. The glass plate B reflects a part of the incident light towards the air film enclosed by the lens L and the plane glass plate G. The reflected beam from the air film is viewed with a microscope. Interference takes place and dark and bright circular fringes are produced. This is due to the interference between the light reflected from the lower surface of the lens and the upper surface of the glass plate G.

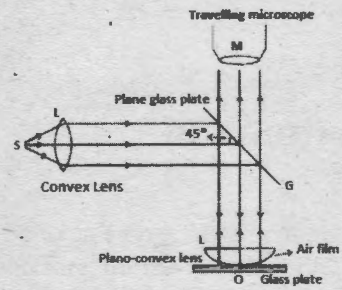


Fig.3.19. Experimental set up for Newton's Rings

Theory of Newton's Rings in Reflected Light

Suppose the radius of curvature of the lens is R and the air film is of thickness t at a distance of r from the point of contact O of the lens and plate combination shown in Fig.3.20(a).

(a) **(a) Interfering rays**

(b) **(b) Diameter of dark and bright rings**

Fig.3.20.

Now if the interference is treated in reflected light through wedge shaped thin film, we have the following condition for bright rings or maxima as:

$$2\mu\, t\cos(r+\theta) = (2n-1)\frac{\lambda}{2} \qquad (53)$$

Where n = 1,2,3............... etc.

As θ is small and for normal incidence, r may also be taken as negligible, therefore

$\cos(r+\theta) = 1$ so

$$2\mu\, t = (2n-1)\frac{\lambda}{2} \qquad (54)$$

Condition for minima or for the dark ring is:

$$2\mu\, t\cos(r+\theta) = n\lambda$$

$$2\mu\, t = n\lambda \qquad (55)$$

Where n = 1,2,3...................................etc. From (54) and (55), it can be shown that for a bright or a dark ring of any particular order, t should be constant. However, in the air film the locus of points having the same thickness is a circle with its centre at the point of contact. Thus, the rings are circular rings having the common centre at the point of contact.

Calculation of Diameters of Dark and Bright Rings

Now using the geometry of the circle in Fig. 3.20(b), we have

$EP \times HE = OE \times (2R\text{-}t)$ (56)

But

$EP = HE = r, OE = PQ = t$

And; $2R\text{-}t = 2R$ (approximately) (57)

$$r^2 = 2Rt$$

$$t = r^2/2R$$

Substituting the values of t in equations (54) and (55):

For bright rings: $r^2 = \frac{(2n-1)\lambda R}{2\mu}$

$$r = \sqrt{\frac{(2n-1)\lambda R}{2\mu}} \quad (58)$$

Or $$D = \sqrt{\frac{(2n-1)\lambda R}{\mu}} \quad (59)$$

For dark rings: $r^2 = \frac{n\lambda R}{\mu}$

But $r = \frac{D}{2}$

Or $$D = \sqrt{\frac{4n\lambda R}{\mu}} \quad (60)$$

Or $$D^2 = \frac{4n\lambda R}{\mu}$$

$$r^2 = n\lambda R$$

$$r = \sqrt{n\lambda R}$$

$$D = \sqrt{2n\lambda R} \quad (61)$$

When n = 0, the radius of the dark ring is zero and the radius of the bright rings is $\sqrt{\frac{\lambda R}{2}}$, therefore the centre is dark. Alternately, dark and bright rings are produced as shown in Fig.3.21.

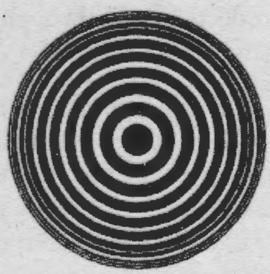


Fig.3.21. Newton's Rings by reflected light

In nutshell we can say that radius of dark ring is proportional to:

(i) $\sqrt{n}$ (ii) $\sqrt{\lambda}$ (iii) $\sqrt{\frac{1}{\mu}}$ (iv) $\sqrt{R}$

Similarly, the radius of the bright ring is proportional to

(i) $\sqrt{\frac{(2n-1)}{2}}$ (ii) $\sqrt{\lambda}$ (iii) $\sqrt{\frac{1}{\mu}}$ (iv) $\sqrt{R}$

If we find the diameter of successive bright or dark rings, it can be seen that spacing between the two successive bright or dark ring goes on decreasing as we proceed from centre to outside. This can be shown as:

If D is the diameter of the dark ring say

$$D = 2r = 2\sqrt{n\lambda R} \qquad (62)$$

For the central dark ring

$$n = 0$$

$$D = 2\sqrt{n\lambda R} = 0$$

This corresponds to the centre of the Newton's rings. While counting the order of the dark rings 1, 2, 3; etc. the central ring is not counted. Therefore, for the first dark ring:

$$n = 1$$

$$D_1 = 2\sqrt{\lambda R}$$

For the second dark ring, $n = 2$,

$$D_2 = 2\sqrt{2\lambda R}$$

and for the n th dark ring,

$$D_n = 2\sqrt{n\lambda R}$$

Take 16[th] and 9[th] rings,

$$D_{16} = 2\sqrt{16\lambda R} = 8\sqrt{\lambda R}$$

$$D_9 = 2\sqrt{9\lambda R} = 6\sqrt{\lambda R}$$

The difference in diameters between the 16 [th] and the 9[th] rings,

$$D_{16} - D_9 = 8\sqrt{\lambda R} - 6\sqrt{\lambda R} = 2\sqrt{\lambda R}$$

Similarly, the difference in the diameters between the fourth and first rings,

$$D_4 - D_1 = 2\sqrt{4\lambda R} - 2\sqrt{\lambda R} = 2\sqrt{\lambda R}$$

Therefore, the fringe width decreases with the order of the fringe and the fringes got closer with increase in their order.

For bright rings,

$$r^2 = \frac{(2n-1)\lambda R}{2} \qquad (63)$$

or

$$D^2 = 2(2n-1)\lambda R \qquad (64)$$

$$r_n = \sqrt{\frac{(2n-1)\lambda R}{2}} \qquad (65)$$

In equation (63), substituting $n = 1, 2, 3$ (number of the ring) the radii of the first, second, third etc. bright rings can be obtained directly.

Newton's Rings by Transmitted Light: In the case of transmitted light, the interference fringes are produced (as shown in Fig. 3.22.) such that for bright rings,

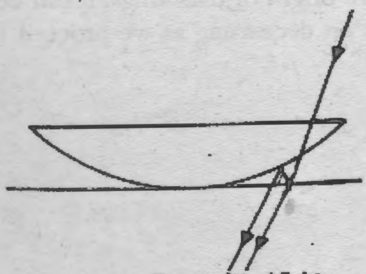


Fig. 3.22. Interference in transmitted region

$$2\mu\, t \cos\theta = n\,\lambda \qquad (66)$$

and for dark rings

$$2\mu\, t \cos\theta = (2n-1)\frac{\lambda}{2} \qquad (67)$$

Here, for air

$$\mu = 1,$$

and $\cos\theta = 1$

For bright rings,

$$2t = n\lambda$$

and for dark rings $2t = (2n-1)\frac{\lambda}{2}$

Taking the value of $t = \frac{r^2}{2R}$, where r is the radius of the ring and R the radius of curvature of the lower surface of the lens, the radius for the bright and dark rings can be calculated.

For bright rings,

$$r^2 = n\lambda R \qquad (68)$$

for dark rings,

$$r^2 = \frac{(2n-1)\lambda R}{2} \qquad (69)$$

where n = 1, 2, 3...... etc.

When, n = 0, for bright rings,

$$r = 0.$$

Therefore, in the case of Newton's rings due to transmitted light as shown in Fig. 3.23, the central ring is bright i.e., just complementary to the one in reflected light. However, the contrast between the bright and dark rings is not as good as that in the reflected light rings.

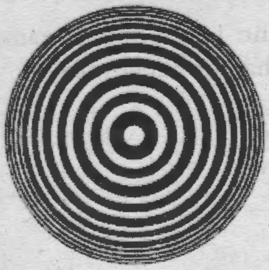


Fig.3.23 Newton's Rings in transmitted light

3.16 Determination of Wavelength of Light Using Newton's Rings Apparatus

Refer to Fig.3.22 again, in the given experimental arrangement; S is an extended source of light placed at the focus of a convex lens. Parallel rays of light from the lens falls on the glass plate B at 45°. The glass plate B reflects a part of the incident light towards the air film enclosed by the lens L and the plane glass plate G. The reflected beam from the air film is viewed with a microscope. The free space between convex side of the lens and the glass plate serves as air filled wedge shaped thin film. The thickness of the air film increases from point of contact of lens to outer edge. Interference takes place and dark and bright circular fringes are produced. This is due to the interference between the light reflected from the lower surface of the lens and the upper surface of the glass plate G.

The interference rings or Newton rings formed can be viewed through the microscope focused on the air film where the rings are formed. After the rings are formed, find diameter of any n^{th} ring and then the diameter of $(n + P)^{th}$ dark ring, with the help of a travelling microscope, where P is well defined. We can then then write the diameters for n^{th} and $n+P$ th dark ring as:

$$D_n^2 = 4n\lambda R \tag{70}$$

$$D_{n+P}^2 = 4(n + P)\lambda R \tag{71}$$

On subtracting the equation (70) from equation (71) and on rearranging we get,

$$\lambda = \frac{D_{n+m}^2 - D_n^2}{4\ P\ R} \tag{72}$$

Using equation (72), λ can be measured.

On looking the equation: $D_n^2 = 4n\mu\lambda R$ (diameter of a dark ring), it seems that merely by putting the values of n, R, λ,μ we can find the wavelength very easily, but here due to the large radius of curvature, the point of contact is not a point and it is a region of several dark fringes. Thus, to locate a fringe of a particular n is not well-defined. Hence, n should be eliminated and we can apply the above mathematical formulation.

e.g. if suppose we have to find the diameter of say 10^{th} and 20^{th} ring then P= 20 − 10=10

Thus, $\lambda = \dfrac{(D_{20})^2 - (D_{10})^2}{4 \times 10R}$

The radius of the curvature of plano-convex lens can be calculated using a spherometer

3.17 Refractive Index of a Liquid Using Newton's Rings

The experiment is performed when there is an air film between the plano-convex lens and the optically plane glass plate. These are kept in a metal container C. The diameter of the n^{th} and the $(n + P)^{th}$ dark rings are determined with air and with the liquid in between the lens and the glass plate, with the help of a travelling microscope(Fig.3.24). The liquid is poured in the container C without disturbing the arrangement. The air film between the lower surface of the lens and the upper surface of the plate is replaced by the liquid. The diameters of the n^{th} ring and the $(n+P)^{th}$ ring are determined.

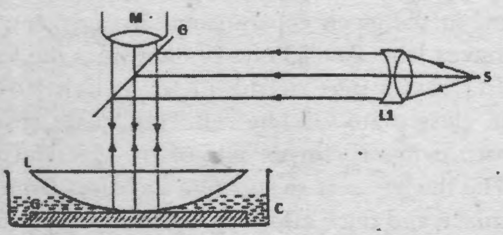


Fig. 3.24. Experimental set up to calculate refractive index of a given liquid

$$D_n^2 = 4n\lambda R \quad (73)$$

$$D_{n+P}^2 = 4(n+P)\lambda R \qquad \text{Without liquid or with air} \quad (74)$$

Substracting (1) from (2), we get:

$$[D_{n+P}^2 - D_n^2]_{air} = 4P\lambda R \quad (75)$$

$$D_n^2 = \frac{4n\lambda R}{\mu} \quad (76)$$

$$D_{n+P}^2 = \frac{4(n+P)\lambda R}{\mu} \qquad \text{With liquid} \quad (77)$$

Substracting (76) from (77) we get:

$$[D_{n+P}^2 - D_n^2]_{liquid} = \frac{4P\lambda R}{\mu} \quad (78)$$

Divide (75) by (77):

$$\mu = \frac{[D_{n+P}^2 - D_n^2]_{air}}{[D_{n+P}^2 - D_n^2]_{liquid}} \quad (79)$$

The relation holds true for bright rings also.

We can also compare the refractive indices of two liquids while doing the experiment with two different liquids. Higher index means contraction of the same order of fringe.

We know, the diameter of n^{th} dark ring with air film is given as:

$$D_{n(air)}^2 = 4n\lambda R \quad (80)$$

The diameter of n^{th} dark ring with liquid film is given as:

$$D_{n(liquid)}^2 = \frac{4n\lambda R}{\mu} \quad (81)$$

from (80) and (81), we get $\frac{D_{n(liquid)}^2}{D_{n(air)}^2} = \frac{1}{\mu}$

or $\frac{D_{n(liquid)}}{D_{n(air)}} = \frac{\text{Diameter of nth ring in liquid}}{\text{Diameter of nth ring in air}} = \frac{1}{\sqrt{\mu}}$

$D_{n(liquid)} < D_{n(air)}$ (as $\mu > 1$)

Note: When a liquid is introduced between the lens and plate, the diameters of the rings decrease or in other words the rings will be contracted.

3.18 Formation of Newton's Rings by Two Curved Surfaces

Newton' ring can be observed if we take the two surfaces curved also. This can be done by replacing the glass plate G in the set up by either using a plano-convex or a plano-concave lens.

(i) Upper Surface Convex and Lower Concave

Take two curved surfaces of radii of curvature R_1 and R_2 in contact at the point O. Here, we have replaced the lower glass plate G with that of a plano-concave lens as shown in the Fig. 3.25. A thin air film is enclosed between the two surfaces. The dark and bright rings are formed and can be viewed with a travelling microscope. Suppose the radius of the n^{th} dark ring = r_n. The thickness of the air film at P, is

$$PQ = PT - QT \tag{82}$$

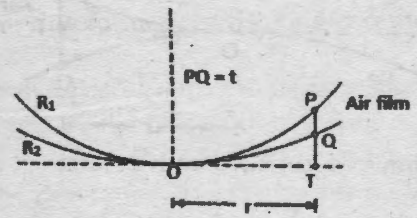


Fig. 3.25 Newton ring formation by two convex surfaces

Here, we have;

$$PT = \frac{r^2}{2R_1}\text{, and } QT = \frac{r^2}{2R_2}$$

$\therefore$ $$PQ = \frac{r^2}{2R_1} - \frac{r^2}{2R_2}$$

But $$PQ = t$$

For reflected light,

$$2\mu t\cos\theta = n\lambda\text{, for dark rings.}$$

Here, for air $$\mu = 1$$

$$\cos\theta = 1$$

$$2t = n\lambda$$

or $$2\left(\frac{r^2}{2R_1} - \frac{r^2}{2R_2}\right) = n\lambda \tag{83}$$

$$r^2\left[\frac{1}{R_1} - \frac{1}{R_2}\right] = n\lambda$$

where $$n = 0,1,2,3 \ldots\text{ etc.}$$

For bright rings,

$$2\mu t\cos\theta = \frac{(2n+1)\lambda}{2}$$

Taking $$\mu = 1$$

and $$\cos\theta = 1$$

$$2t = \frac{(2n+1)\lambda}{2}$$

or $$r^2\left[\frac{1}{R_1} - \frac{1}{R_2}\right] = \frac{(2n+1)\lambda}{2} \tag{84}$$

where $$n = 0,1,2,3 \ldots\text{ etc.}$$

For the 10^{th} bright ring, the value of $n = 10 - 1 = 9$

$\therefore$ For n the bright ring,

$$r_n^2\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{[2(n-1)+1]\lambda}{2} = \frac{(2n-1)\lambda}{2} \tag{85}$$

(ii) Both Upper and Lower Surface Convex

Take two curved surfaces of radii of curvature R_1 and R_2 in contact at the point O. Here, we have replaced the lower glass plate G with that of a plano-convex lens as shown in the Fig. 3.26. A thin air film is enclosed between the two surfaces. The dark and bright rings are formed and can be viewed with a travelling microscope.

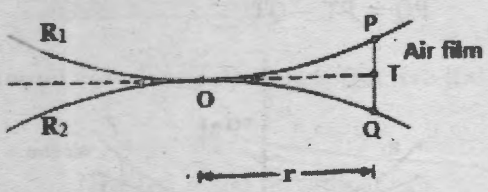


Fig. 3.26. Newton rings formation by a concavo-convex lens

$$PQ = PT + QT$$
$$= \frac{r^2}{2R_1} + \frac{r^2}{2R_1} \qquad (86)$$

For dark rings, $2PQ = n\lambda$

$$r^2\left[\frac{1}{R_1} + \frac{1}{R_2}\right] = n\lambda \qquad (87)$$

For bright rings, $2PQ = (2n + 1)\frac{\lambda}{2}$

$$r^2\left[\frac{1}{R_1} + \frac{1}{R_2}\right] = (2n + 1)\frac{\lambda}{2}$$

when $n = 0,1,2,3 \ldots$ etc.

For the first bright ring, $n = 0$

$$\therefore \quad r_1^2\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{\lambda}{2}$$

For the 10th bright ring, $n = 9$

$$r_{10}^2\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = [2(n) + 1]\frac{\lambda}{2}$$

For the nth bright ring

$$r_n^2\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = [2(n-1) + 1]\frac{\lambda}{2}$$

$$r_n^2\left[\frac{1}{R_1} + \frac{1}{R_2}\right] = \frac{(2n-1)\lambda}{2} \qquad (88)$$

3.19 Newton's Rings with a Bright Centre due to Reflected Light

Whenever there is an air film between the lens and the plane glass plate, the rings formed by reflected light have a dark centre. We know that the two surfaces are just in contact at the centre and the two interfering rays are reflected under different conditions. As a result, a path difference of half a wave length occurs because of one of the rays undergoing a phase change of π, when reflected from the glass plate.

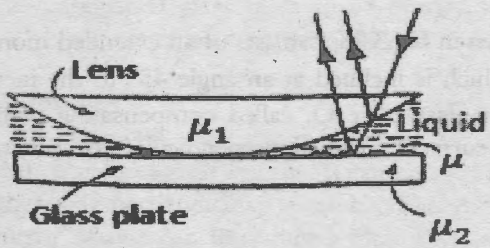


Fig. 3.27. Newton ring formation with bright centre in reflected light

Let us now introduce a transparent liquid of refractive index μ between the two surfaces in contact. If the refractive index of the material of the lens and glass plate is μ_1 and μ_2 respectively such that $\mu_1 > \mu > \mu_2$ as shown in Fig.3.27. If we put a little oil of sassafras between a convex lens of crown glass and a plate of flint glass, the above condition can be met easily. The reflections in both the cases will be from denser to rarer medium and the two interfering rays are reflected under the similar conditions. Therefore, in this case the central spot will be bright. Even if $\mu_1 < \mu < \mu_2$, the central spot will be bright, because a path difference of $\frac{\lambda}{2}$ takes place at both the upper and the lower glass-liquid surfaces. The two interfering beams are reflected under similar conditions and the central spot is bright due to reflected light.

3.20 Newton's Rings with White Light

If white light is used, the central dark spot remains the same in reflected light. However, since the diameter of the rings be it bright or dark depends upon the wavelength of light used, so when white light is used, the diameter of the rings of the different colours will be different and coloured rings are observed. Only the first few rings are clear and gradually after that due to overlapping of the rings of different colours, the rings cannot be viewed as we move away from the centre. This overlapping increase to so much of extent that the rings cannot be distinguishably seen resulting in a uniform illumination.

3.21 Replacing Glass Plate by Plane Mirror

If the glass plate P is replaced by a plane mirror in case of Newton's rings, then the rings become invisible. This is because the light is completely reflected from the plane mirror. The amplitude of the light wave reflected by the plane mirror will be much greater than the amplitude of the light wave reflected by convex surface of the lens. Due to this difference in the amplitudes (the two superimposing waves must have equal amplitudes for a good contrast) there will be no observable contrast between bright and dark fringes thereby showing invisibility of rings.

3.22 Michelson Interferometer

It is one of the most prominent optical instruments, which is used for doing various fine measurement regarding lengths using the phenomena of interference by division of amplitude. It was designed and developed by Albert Abraham Michelson for the measurement of small lengths and standardization of meter by the application of interference phenomena, and so is referred to as Michelson Interferometer.

Principle: Interference by division of amplitude is the principle behind the development of this optical instrument.

Construction:

Michelson interferometer, shown in Fig.3.28, consists of an extended monochromatic source of light S, a semi-silvered glass plate P which is inclined at an angle 45° to the incident beam of light. There is another unpolished transparent glass plate Q, called compensating plate. Two mirrors M_1 and M_2 highly silvered on their front surfaces inclined perpendicular to each other. T is the telescope for observing interference pattern.

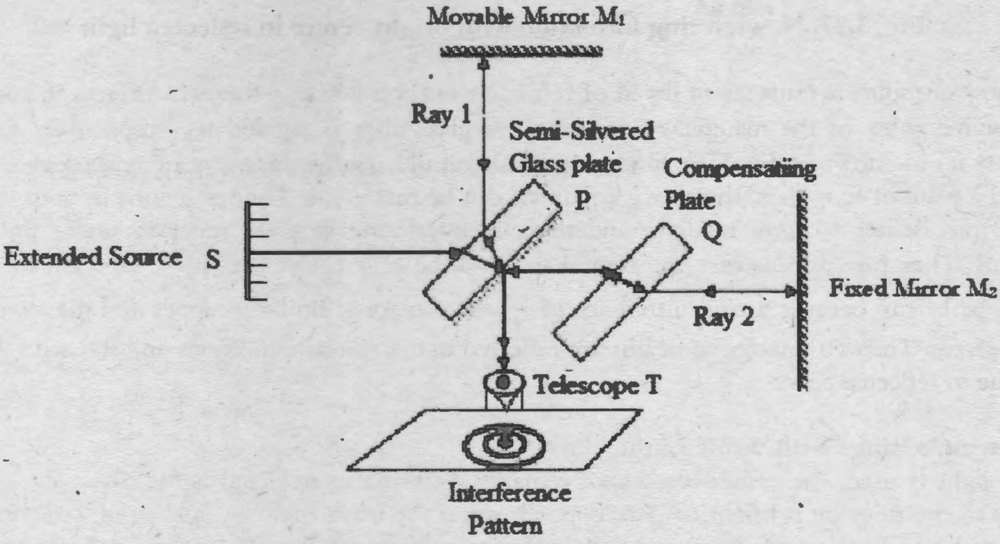


Fig. 3.28 Michelson Interferometer

Working

When a light beam from a source S, falls at the semi-silvered plate P, the amplitude of the light beam divides into two equal amplitudes, as reflected and transmitted beams 1 and 2. The reflected beam 1, moves towards M_1 and reflected back to P and the transmitted beam 2, moves towards M_2 and is reflected back from mirror M_2 to plate P. These two reflected 1 and 2, will interfere, constructively or destructively depending on their path/phase difference. The fringe pattern thus will be obtained and can be observed in the telescope T.

Actually, the two waves, which are generated as a result of division of amplitude at the surface of the semi-silvered glass plate P, can be considered as the originated from two virtual coherent sources, M_1 and M_2. These two reflected waves meet on the screen of the telescope and the interference pattern can be seen the telescope.

Role of Compensating Plate Q

The plate P is semi-silvered at its back surface. So, the incident wave from the source S, is divided into two parts at the back surface. The wave 1 reflected towards the mirror M_1 and the wave 2 is transmitted towards the mirror M_2. Now the wave 1, reaches at the telescope after, reflected from M_1 and transmitted through the plate P. The second wave 2 also reaches at the telescope after reflected from M_2 and the reflected from the back surface of the plate P. Now if the physical path lengths of the two mirrors are exactly same, even then there will be a path difference between the two. It is because that after division at the semi-silvered back surface of the plate P, the wave 1 is passed twice, through the plate P but the wave 2 travelled in the air only. So, a path is created between the two. To compensate this path, difference a second unpolished glass plate Q of exactly same dimensions, is introduced between the paths of wave 2. This plate is called compensating plate.

Fringe Pattern and Shape of Fringes

When we see through the telescope, the two waves seem to be coming out from two mirrors overlapping on each other. The virtual image of mirror M_2 can be seen parallelly placed near the mirror M_1, if the two mirrors are physically perpendicular to each other. As the mirror M_1 is movable and tiltable, so by moving the mirror M_1 we can introduce a fine physical path difference between two mirrors. This path difference between M_1 and M_2 can be considered as the parallel air thin film at the place of mirror M_1. Now if we tilt the mirror M_1, then a wedge-shaped air thin film will be formed at mirror M_1. The interference pattern and the shape of the fringes obtained in both the cases, can be easily explained, using the well-versed theory of interference through parallel and wedge-shaped thin films (Fig. 3.29)

Types of fringes

1. **Circular fringes:** When mirror M_2 is exactly perpendicular to mirror M_1 or the mirror M_1 and the virtual mirror M_2' (image of M_2) are parallel, an air film of constant thickness is enclosed between them and circular fringes are obtained.
2. **Localised fringes:** When mirror M_2 is not exactly perpendicular to mirror M_1 or the mirror M_1 and the virtual mirror M_2' (image of M_2) are inclined, the air film enclosed between them is wedge shaped and fringes appear to be straight. If one of the mirrors is moved, the fringes move across the field.

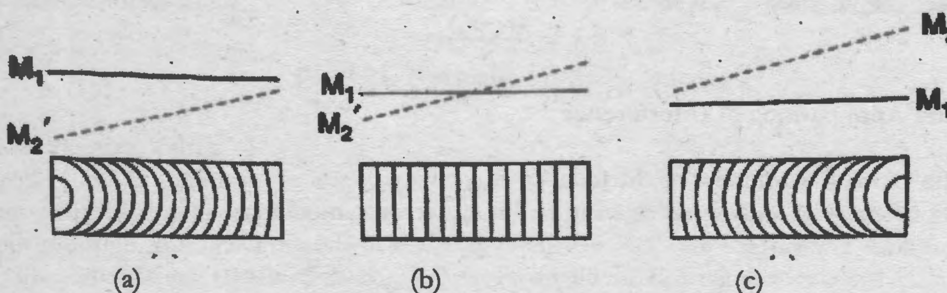


Fig. 3.29 Localised Fringes

Fig.3.29 above shows when the two mirrors M_1 and M_2' intersect in the middle straight fringes are observed as in Fig.3.29 (b). When the two mirrors are inclined as shown in Fig.3.29 (a) and (c), curved fringes with convexity towards the thin edge of the wedge are observed.

Applications of Michelson Interferometer

Michelson interferometer is widely being used for various measurement purposes. Thin film thickness, refractive index, standardization of meter, laser interferometry etc. are some of the applications.

1. Thin Film Thickness and Refractive Index Measurement: Michelson interferometer is set for zero path difference and the pattern is obtained in the telescope. As a thin film is introduced in the path of ray 1, the fringe pattern is displaced in the field of view of the telescope. Now with the help of micrometre, the mirror is move in such a way the pattern comes to its original position. Let x be the distance moved by the mirror on the micrometre scale. Using following equation, the thickness of the film can be calculated.

$$2x = 2(\mu - 1)t \quad \text{or} \quad t = \frac{x}{(\mu-1)}$$

Where μ is the refractive index of the material and t is the thickness of the film. If we know μ, t can be calculated or vice versa.

2. Standardization of Meter: By counting the number of shifted fringes or number of wavelengths, when a standard scale of length is introduced between one arm of the interferometer, one can standardize that length, as the particular number of wavelengths of a particular colour, equivalent to a particular length never change. So, using that numbers of wavelengths, a standard meter can be reproduced, without any error in it.

3. Determination of Wavelength of Monochromatic Light: Michelson Interferometer is first set for circular fringes with central bright spot. If t be the thickness of the air film enclosed between the two mirrors and n be the order of the spot obtained, then as for normal incidence cos r = 1, we have:

$$2t + \frac{\lambda}{2} = n\lambda$$

If M_1 is moved $\lambda/2$ away from M_2', then an additional path difference of λ will be introduced and hence $(n+1)^{th}$ bright spot appears at the centre of field. Thus, each time when M_1 moves through a distance $\lambda/2$ next bright spot appears at the centre of the field. Let N ne the number of fringes that cross the centre of field when the mirror M_1 is moved from initial position x_1 to the final position x_2, then

$$N\frac{\lambda}{2} = x_2 - x_1$$

Or

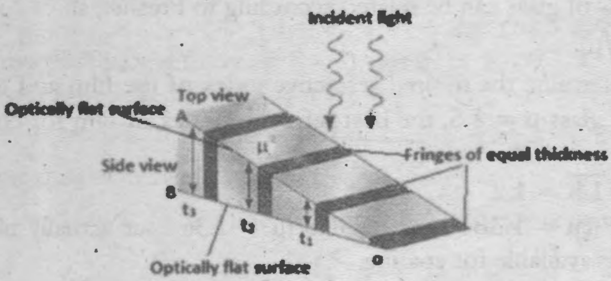
$$\lambda = \frac{(x_2 - x_1)}{N}$$

3.23 General Applications of Interference

(1) Test The Optical Planeness of Surface

The optical planeness of a given surface can be known if we can compare it with another optically flat or a plane surface. The surface must be sufficiently finished to show reflection of light but the ordinary ground surfaces are too irregular to show light wave interference bands. If we take these two surfaces inclined to each other, we get a thin wedge-shaped film with gradually increasing thickness. In order to give a definite value to the wavelength, the surface under the test is illuminated with monochromatic source of light.

Starting from the point of contact say o of the two plane surfaces OA and OB and moving to the broader part of the film as in Fig. 3.30. When, $2t_1 = \lambda$ (for normal incidence) interference occurs causing the first dark fringe which is straight line parallel to the edge of the wedge. When now $2t_2 = 3/2\lambda$, the wave trains will superimpose on each other and a bright fringe will be produced which is again parallel to the edge of the wedge. Moving further, when again $2t_3 = 2\lambda$, second dark fringe appears and so on. Hence, if the surfaces are plane the fringes will be straight lines, equally spaced and parallel to the line of intersection of the two surfaces shown in Fig. 3.30. The fringes observed are of equal thickness because each fringe is the locus of the points at which the thickness of the film has a constant value. It is an important application of the phenomenon of interference. If the fringes are not of equal thickness, it means the surfaces are not plane shown in Fig. 3.31



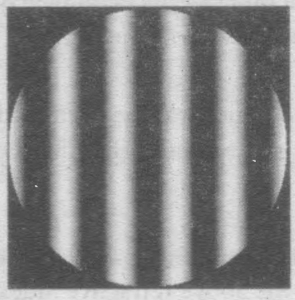
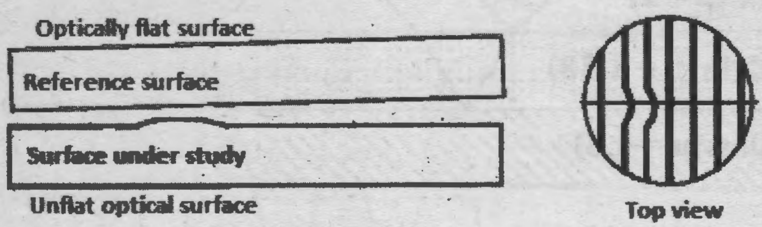


Fig.3. 30. Optical planeness of a surface



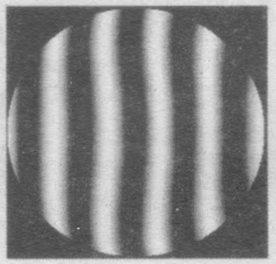


Fig.3.31. Optically un-flat surface

(2) Non-Reflecting Films

It is one of the most important practical applications of interference. The principle of thin film can be applied to reduce the reflection from a lens or a prism to its minimum by coating the surface under study with a thin non-reflecting film. When a beam of light is incident normally on a glass, nearly 4% of incident light is reflected while remaining 96% is transmitted. This loss of light by reflection is undesirable and has to be eliminated in camera lenses and in many other optical instruments, thus one can use a non- reflecting glass. Similarly, the lens-surfaces are often coated with $\frac{\lambda}{4\mu}$ thick non-reflecting film, the refractive index of film being less than that of lens. e.g., A glass lens $(\mu = 1.5)$ can be coated with MgF_2 film $(\mu = 1.38)$ shown in Fig. 3.32 and the thickness of the film should be adjusted so as to obtain minima in reflected light. When light is incident normally from air on glass through MgF_2, the reflections at air – MgF_2 film and glass interfaces are such that, the two phases (each of π) cancel out i.e. the optical path difference for destructive interference is $(2n-1)\frac{\lambda}{2}$ which leads to:

$$2\mu_f t = (2n-1)\frac{\lambda}{2}, \qquad n = 1,2,3, \ldots \ldots \tag{89}$$

For the calculation of minimum thickness, put $n = 1$ i.e.

$$2\mu_f t_{min} = \frac{\lambda}{2}$$

$$t_{min} = \frac{\lambda}{4\mu_1} \tag{90}$$

To annul each other's effect completely, the two reflected beams should be of equal intensity.
The refractive index μ_f of the film and μ_g of glass can be related according to Fresnel, as

$$\mu_f = \sqrt{\mu_g}$$

The above condition can be used to determine the desired refractive index of the film and the film is referred to as an anti-reflecting film. For glass $\mu = 1.5$, the best refractive index of film for coating will be

$$\mu_f = \sqrt{1.5} = 1.2$$

The materials commonly used are $MgF_2(\mu = 1.38)$ and cryolite ($\mu = 1.36$) but actually no durable material with the above refractive index is available for coating.

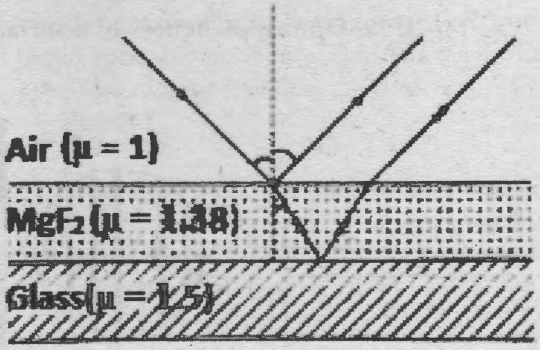


Fig. 3.32. Non-reflecting film

******Solved Examples*******

Based on YDSE

Ex.1 Two coherent sources whose intensity ratio is 81:1 produce interference fringes. Deduce the ratio of maximum intensity to minimum intensity in fringe system.

Sol. We have

$$I_{max} = (a_1 + a_2)^2$$
$$I_{min} = (a_1 - a_2)^2$$
$$\frac{I_{max}}{I_{min}} = \frac{(a_1+a_2)^2}{(a_1-a_2)^2} \tag{1}$$

Given

$$\frac{I_1}{I_2} = \frac{a_1^2}{a_2^2} = \frac{81}{1}$$

or

$$\frac{a_1}{a_2} = \frac{9}{1} \Rightarrow a_1 = 9a_2$$

Substituting this value of a_1 in equation (1), we get

$$\frac{I_{max}}{I_{min}} = \left(\frac{9a_1+a_2}{9a_1-a_2}\right)^2 = \left(\frac{10}{8}\right)^2 = \frac{100}{64} = \frac{25}{16}$$
$$I_{max}: I_{min} = 25: 16$$

Ex.2 Two identical coherent waves produce interference pattern. Find the ratio of intensity at the centre of a bright fringe to the intensity at a point one quarter of the distance between two fringes from the centre.

Sol. The intensity I at any point is given by

$$I = a_1^2 + a_2^2 + 2a_1a_2\cos\delta$$

As the coherent waves are identical the amplitudes of two waves are equal i.e. $a_1 = a_2 = a$; then the intensity equation (1) becomes

$$I = a^2 + a^2 + 2a^2\cos\delta = 2a^2(1 + \cos\delta)$$

At the centre the phase difference, $\delta = 0, \cos 0^\circ = 1$

$\therefore$ Intensity at centre, $I_0 = 2a^2(1 + 1) = 4a^2$

The phase difference between two consecutive fringes is 2π, therefore the phase difference at a distance one quarter between two fringes will be $\frac{2\pi}{4} = \frac{\pi}{2}$.

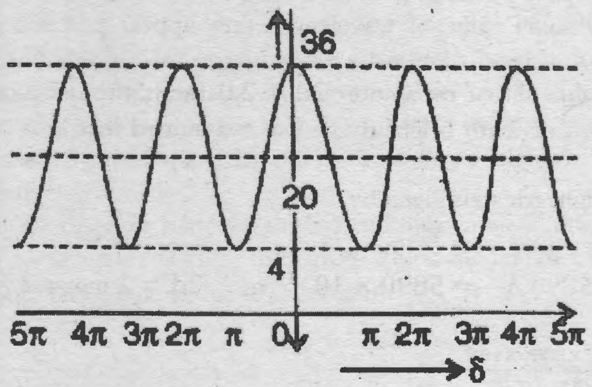
If I_1 is the intensity at a point one quarter of the distance between two fringes from the centre, then

$$I_1 = 2a^2\left[1 + \cos\frac{\pi}{2}\right] = 2a^2$$

$\therefore$ Required ratio, $\frac{I_0}{I_1} = \frac{4a^2}{2a^2} = 2$

Ex.3. Two waves of amplitudes 4 and 2 units are superposed with their vibrations parallel. Deduce the ratio of the maximum to minimum intensity as phase relation varies. Also plot a curve of intensity versus phase difference.

Sol.



The maximum amplitude $A_{max} = 4 + 2 = 6$ units.

The minimum amplitude $A_{min} = 4 - 2 = 2$ units.

$$\frac{I_{max}}{I_{min}} = \frac{A_{max}^2}{A_{min}^2} = \frac{6^2}{2^2} = \frac{36}{4} = 9$$

For plotting the curve, we have

$$I = a_1^2 + a_2^2 + 2a_1a_2\cos\delta = 4^2 + 2^2 + 2 \times 4 \times 2\cos\delta$$
$$= 16 + 4 + 16\cos\delta$$
$$= 20 + 16\cos\delta$$

Fig. above represents the variation of $16\cos\delta$ over a background of 20 units.

Ex.4 Two coherent sources of intensity ratio β interfere. Prove that in intensity pattern

$$\frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{2\sqrt{\beta}}{1 + \beta}$$

Sol. We have $I_{max} = (a_1 + a_2)^2$

$I_{min} = (a_1 - a_2)^2$

Intensity of first wave, $I_1 = a_1^2$

Intensity of second wave, $I_2 = a_2^2$.

i.e., $\frac{I_1}{I_2} = \frac{a_1^2}{a_2^2} = \beta$

or, $\frac{a_1}{a_2} = \sqrt{\left(\frac{I_1}{I_2}\right)} = \sqrt{\beta}$

$$\frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{(a_1+a_2)^2 - (a_1-a_2)^2}{(a_1+a_2)^2 + (a_1-a_2)^2}$$

$$= \frac{4a_1a_2}{2(a_1^2+a_2^2)}$$

$$= \frac{2a_1a_2}{a_2^2\left(\frac{a_1^2}{a_2^2}+1\right)}$$

$$= \frac{2(a_1/a_2)}{\frac{a_1^2}{a_2^2}+1} = \frac{2\sqrt{\beta}}{\beta+1}$$

i.e. $$\frac{I_{max}-I_{min}}{I_{max}+I_{min}} = \frac{2\sqrt{\beta}}{1+\beta}$$

Ex.5 Two straight and narrow parallel slits 3 mm apart are illuminated by a monochromatic light of wavelength 5900 Å. Fringes are obtained on a screen 60 cm away from the slits. Calculate the fringe width. Comment on the shape of fringes.

Sol. Distance between slits, $2d = 3mm = 3 \times 10^{-3}$m ; Distance of the screen from the slits, $D = 60$ cm $= 60 \times 10^{-2}$ m $= 0.60$ m ; Wavelength of light, $\lambda = 5900$ Å$= 5900 \times 10^{-10}$ m $= 5.9 \times 10^{-7}$m

$\therefore$ $$\text{fringe width, } \omega = \frac{D\lambda}{2d} = \frac{0.60 \times 5.9 \times 10^{-7}}{3 \times 10^{-3}}$$

$$= 1.18 \times 10^{-4} \text{ m} = 0.118 \text{ mm}$$

The fringes are hyperboloids of revolution in the space between parallel-slits and the screen. On the screen they are hyperbolae, but due to extremely small value of wavelength they appear practically straight lines.

Ex.6 Sodium light ($\lambda = 5890$ Å) falls on a double slit of separation $2d = 2.0$ mm , the distance between slits and screen is 4 cm. Locate the position of tenth bright fringe if it is assumed that $D >> 2d$.

Sol. If we assume $D >> 2d$, the position of n^{th} bright fringe is given by

$$y_n = \frac{nD\lambda}{2d}$$

Here $n = 10, D = 4$ cm $= 4 \times 10^{-2}$ m, $\lambda = 5890$ Å $= 5890 \times 10^{-10}$ m , $2d = 2$ mm$= 2 \times 10^{-3}$ m

$\therefore$ Position of tenth bright fringe, $$y_{10} = \frac{10\times4\times10^{-2}\times5890\times10^{-10}}{2\times10^{-3}}$$

$$= 1.178 \times 10^{-4}\text{m} = 0.1178\text{mm.}$$

Ex.7 In a Young's double slit experiment the angular width of a fringe formed on a distant screen is 0.1°. The wavelength of light used is 6000 Å. What is the spacing between the slits?

Sol. Angular fringe width $\omega_\theta = \frac{\lambda}{2d}$

Spacing between slits $2d = \frac{\lambda}{\omega_\theta}$

Here $\lambda = 6000$ Å $= 6 \times 10^{-7}$ m , $\omega_\theta = 0.1° = \frac{0.1 \times \pi}{180}$ radians

$$\therefore\ 2d = \frac{6 \times 10^{-7}}{\left(0.1 \times \frac{3.14}{180}\right)} = \frac{6 \times 10^{-7} \times 180}{0.1 \times 3.14} \text{m} = 3.44 \times 10^{-3} = 0.344 \text{ nm}$$

Based on Fresnel's biprism

Ex.8 In a Fresnel's biprism experiment the fringe width observed is 0.087 mm. What will it become if the slit to biprism is reduced to $\frac{3}{4}$ of the original distance (all else remaining unchanged).

Sol. The separation between virtual sources in biprism experiment is

$$2d = 2a(\mu - 1)\alpha$$

Where a = distance between slit and the biprism

α = refracting angle of the biprism

If new distance between slit and biprism is a′

then $a' = \frac{3}{4}a$

∴ New separation between coherent sources.

$$2d' = 2.\left(\frac{3}{4}a\right)(\mu - 1)\alpha \quad (2)$$

Dividing (2) by (1), we get

$$\frac{2d'}{2d} = \frac{3}{4}$$

Fringe width $\omega = \frac{D\lambda}{2d}$

$$\therefore \quad \frac{\omega'}{\omega} = \frac{2d'}{2d} = \frac{4}{3} \quad \text{(as D, } \lambda \text{ are unchanged)}$$

∴ New fringe width $\omega' = \frac{4}{3} \times \omega$

$$= \frac{4}{3} \times 0.087 \text{ mm} = 0.116 \text{ mm}$$

Ex.9 Fringes are produced by a Fresnel's biprism in the focal plane of a reading microscope which is 100 cm from the slit. A lens inserted between the biprism and the eye-piece gives two images of the slit in two positions of the lens. In one case the two images of the slit are 4.05 mm apart and in the other case 2.90 mm apart. If sodium light of wavelength 5893 Å is used, find the width of the interference fringes.

If the distance between the slit and biprism is 10 cm and refractive index of the material of the biprism 1.5, calculate the angle in degrees which the inclined faces of the biprism make with its base.

Sol. The fringe width ω is given by

$$\omega = \frac{D\lambda}{2d}$$

We have $2d = \sqrt{(d_1 d_2)}$

where d_1 and d_2 are the distances between the two images for two different positions of the lens between biprism and eye-piece.

Here $d_1 = 4.05 \text{ mm} = 0.405 \text{ cm}, d_2 = 2.980 \text{ mm} = 0.290 \text{ cm}$

$$2d = \sqrt{(0.405 \times 0.290)} = 0.342 \text{ cm}$$

Also given D = 100 cm and $\lambda = 5893 \times 10^{-8}$ cm

$$\omega = \frac{D\lambda}{2d} = \frac{100 \times 5893 \times 10^{-8}}{0.342} = 0.0172 \text{ cm}$$

In terms of refracting angle (α) of the biprism, the separation (2d) of two coherent sources is given by

$$2d = 2a(\mu - 1)\alpha$$

where a is the distance between slit and biprism.

Here $a = 10$ cm, $\mu = 1.5, 2d = 0.342$ cm

$$\alpha = \frac{2d}{2a(\mu - 1)}$$

$$\alpha = \frac{0.342}{2 \times 10(1.5 - 1)}$$

$$= 0.0342 \text{ radians}$$
$$= 0.0342 \times \frac{180}{\pi} \text{ degrees}$$
$$\approx 2°.$$

Ex.10 In an experiment with Fresnel's biprism, bands 0.0196 cm in width are observed at a distance 100 cm from the slit. A convex lens is then put between the biprism and the eye-piece so as to give an image of the sources at a distance of 100 cm from the slit. The distance apart of the images is found to be 0.70 cm the lens being 30 cm from the slit. Calculate the wavelength of light used.

Sol. Given fringe width $\omega = 0.196$ mm $= 0.0196$ cm, $D = 100$ cm

When the lens is at a distance of 30 cm from the slit, the distance apart of the images is 0.70 cm, we have

$u = 30$ cm, $v = (100 - 30) = 70$cm, $d_1 = 0.70$ cm

From magnification formula, we have

$$\frac{d_1}{2d} = \frac{v}{u}$$
$$\frac{0.70}{2d} = \frac{70}{30}$$
$$2d = \frac{30 \times 0.70}{70} = 0.30 \text{ cm}$$

∴ Wavelength of light used,

$$\lambda = \frac{2d}{D}\omega$$
$$= \frac{0.30 \times 0.0196}{100} = 5880 \times 10^{-8} \text{ cm} = 5880 \text{ Å}$$

Ex.11 A Fresnel biprism arrangement is set with sodium light $(\lambda = 5893 \text{ Å})$ and in the field of view of the eye-piece we get 62 fringes. How many fringes shall we get if we replace the source by mercury lamp using (a) green filter $(\lambda = 5461 \text{ Å})$, (b) violet filter $(\lambda = 4358 \text{ Å})$.

Sol. The fringe width, $\omega = \frac{D\lambda}{2d}$ (for sodium light)

As D and 2d are constants in this problem, the change in λ will cause a change in ω.

(a) If the source is replaced by mercury light (wave-length λ_1) the fringe width ω_1 is given by

$$\omega_1 = \frac{D}{2d}\lambda_1$$

Dividing (1) by (2), we get

$$\frac{\omega}{\omega_1} = \frac{\lambda}{\lambda_1}$$

The width of the field of view = $62\,\omega$ (given)

Let the field of view contain n_1 fringes when we used green filter.

Then width of field of view = $n_1\omega_1$

$$n_1\omega_1 = 62\,\omega$$
$$n_1 = 62\frac{\omega}{\omega_1}$$

Substituting value of $\frac{\omega}{\omega_1}$ from (3) in (6), we get

$$n_1 = 62\frac{\lambda}{\lambda_1} = 62 \times \frac{5893 \times 10^{-10}}{5461 \times 10^{-10}} = 67 \text{ (whole number).}$$

(b) Let the field of view contain n_2 fringes when we use violet filter (wave-length λ_2). We have as above

$n_2 = 62.\frac{\lambda}{\lambda_2} = 62 \times \frac{5893 \times 10^{-10}}{4358 \times 10^{-10}} = 84$ (whole number).

Ex.12 In a biprism experiment with sodium light $(\lambda = 5893\ \text{Å})$ the micrometer reading is 2.32 mm. when the eye-piece is placed at a distance of 100 cm from the source. If the distance between two virtual sources is 2 cm. Find the new reading of the micrometer if the eye-piece is moved such that 20 fringes cross the field of view.

Sol. Given $\lambda = 5893\ \text{Å} = 5893 \times 10^{-10}$ m

$D = 100$ cm $= 1$m , $2d = 2$ cm$= 2 \times 10^{-2}$ m

Fringe width
$$\omega = \frac{D\lambda}{2d}$$
$$= \frac{1.00 \times 5893 \times 10^{-10}}{2 \times 10^{-2}}$$
$$= 2.9465 \times 10^{-5}\ \text{m}$$

∴ The distance moved for 20 fringes to cross the field of view
$$= 20\ \omega$$
$$= 20 \times 2.9465 \times 10^{-5}$$
$$= 5.893 \times 10^{-4}\text{m} = 0.5893\ \text{mm} = 0.59\ \text{mm}$$

∴ The initial reading of micrometer $= 2.32$ mm

∴ Final reading of the micrometer $= (2.32 \pm 0.59)$ mm
$$= 2.91\ \text{mm}\ \ \text{or}\ \ 1.73\ \text{mm}$$

Ex.13 Interference fringes are obtained using a glass biprism of base angle 2° each and refractive index 1.50. If the distance between the slit and the biprism is 10 cm. Calculate the fringe width in the interference pattern obtained at distance of 90 cm away from the biprism, when the slit is illuminated by light of wavelength (i) 5890 Å, (ii) 4600 Å.

Sol. Distance of the slit source from biprism, a = 10 cm

Distance of the screen from biprism, b = 90 cm.

$\mu = 1.5$, angle of prism, $\alpha = 2° = \frac{2\pi}{180}$ radian

∴
$$D = a + b = 10 + 90 = 100\text{cm} = 1\text{m}$$

and
$$2d = 2a(\mu - 1)\alpha = 2 \times 10 \times 10^{-2}(1.5 - 1)\frac{2\pi}{180} = \frac{\pi}{9} \times (10^{-2})$$

The fringe width,
$$\omega = \frac{D\lambda}{2d}$$

(i) When $\lambda = 5890\ \text{Å} = 5890 \times 10^{-10}$m
$$\omega = \frac{1.00 \times 5890 \times 10^{-10}}{\left(\frac{\pi}{9} \times 10^{-2}\right)}$$
$$= \frac{1.00 \times 5890 \times 10^{-10} \times 9}{(\pi \times 10^{-2})}$$
$$= 1.688 \times 10^{-4}\text{m}$$

(ii) When $\lambda = 4600\text{Å} = 4600 \times 10^{-10}$ m
$$\omega = \frac{1.00 \times 4600 \times 10^{-10}}{\left(\frac{\pi}{9} \times 10^{-2}\right)}$$
$$= \frac{1.00 \times 4600 \times 10^{-10} \times 9}{(\pi \times 10^{-2})}$$
$$= 1.318 \times 10^{-4}\ \text{m}$$

Ex.14 In a certain region of interference we get 492th order maximum for sodium 5889.97 Å line. What will be the order of interference at the same place for (i) Sodium 5895.93 Å, (ii) Mercury 5460.7 Å and (iii) Hydrogen 4861.4 Å?

Sol. The nth order maximum is obtained when the path difference

$$\Delta = n\lambda$$

∴ Given $n = 492$ and $\lambda = 5889.97\ \text{Å} = 5889.97 \times 10^{-8}$ cm

∴ At given point P the path difference,

$$\Delta = 492 \times 5889.97 \times 10^{-8} \text{ cm}$$
$$= 2.890 \times 10^{-2} \text{ cm}$$

(i) If for sodium light $\lambda_1 = 5895.93$ Å, the order of interference at P is n_1th then

$$\Delta = n_1\lambda_1 = 2.898 \times 10^{-2}$$

or $$n_1 = \frac{2.898 \times 10^{-2}}{\lambda_1} = \frac{2.898 \times 10^{-2}}{5893.93 \times 10^{-8}} = 491 \text{ (whole number)}$$

(ii) If for mercury light $\lambda_2 = 5460.7$Å the order of interference at P is n_2 th then

$$\Delta = n_2\lambda_2 = 2.898 \times 10^{-2}$$

or $$n_2 = \frac{2.898 \times 10^{-2}}{\lambda_2} = \frac{2.898 \times 10^{-2}}{5460.7 \times 10^{-8}} = 531 \text{ (whole number)}$$

(iii) If for Hydrogen light $\lambda_3 = 4861.4$ Å ,the order of interference at P is n_3th then

$$\Delta = n_3\lambda_3 = 2.898 \times 10^{-2}$$

or $$n_2 = \frac{2.898 \times 10^{-2}}{\lambda_3} = \frac{2.898 \times 10^{-2}}{4861.4 \times 10^{-8}} \approx 596$$

Based on shifting of fringes

Ex.15 When a thin piece of glass 3.4×10^{-4} cm thick is placed in the path of one of the interfering beams in a biprism experiment, it is found that the central bright fringe shifts through a distance equal to width of four fringes. Find the refractive index of the piece of glass. (Wavelength of light = 5893 Å).

Sol. Displacement of fringe $\Delta = \frac{\omega}{\lambda}(\mu - 1)t$

Given $\Delta y = 4\omega, t = 3.4 \times 10^{4}\text{cm} = 3.4 \times 10^{-6}\text{m}, \lambda = 5893\text{Å} = 5893 \times 10^{-10}$ m

∴ $$4\omega = \frac{\omega}{\lambda}(\mu - 1)t$$

⇒ $$(\mu - 1)t = 4\lambda \quad \text{or} \quad \mu - 1 = \frac{4\lambda}{t}$$

⇒ $$\mu = 1 + \frac{4\lambda}{t}$$
$$= 1 + \frac{4 \times 5893 \times 10^{-10}}{3.4 \times 10^{-6}} = 1 + 0.69 = 1.69$$

Ex. 16 A biprism forms interference fringes with monochromatic light of wavelength 5450 Å. On introducing a thin glass plate ($\mu = 1.5$) in the path of one of the interfering beams the central bright band shifts to the position previously occupied by third bright fringe. Giving reasons state whether the fringe width changes and find the thickness of the plate.

Sol. The introduction of a thin plate in the path of one of the interfering beams displaces the entire pattern through a distance $\frac{\omega}{\lambda}(\mu - 1)t$ [ω is fringe width, μ is the refractive index of the plate of thickness t, λ is the wavelength of monochromatic light] towards the side on which the plate is placed and hence there is no change in the fringe width.

When central bright fringe is shifted to nth bright fringe, we have $\frac{\omega}{\lambda}(\mu - 1)t = n\omega$

$$t = \frac{n\lambda}{\mu - 1}$$

Here $\mu = 1.5, n = 3, \lambda = 5450Å = 5450 \times 10^{-10}$ m

$$t = \frac{3 \times 5450 \times 10^{-10}}{1.5 - 1}$$

$$= \frac{3 \times 5450 \times 10^{-10}}{0.5} = 3.27 \times 10^{-6} = 3.27\ \mu m$$

Ex.17 A two slit Young's interference experiment is done with monochromatic light of wavelength 6400 Å. These slits are 2 mm apart and the fringes are observed on a screen placed 10 cm away from the slits and it is found that the interference pattern shifts by 5 mm, when a transparent plate of thickness 0.5 mm is introduced in the path of one of the slits. What is the refractive index of the transparent plate?

Sol. Here $2d = 2\text{ mm} = 2 \times 10^{-3}\text{m}, D = 10\text{cm} = 0.10$ m

$$\lambda = 6000Å = 6000 \times 10^{-10}\text{m} = 6 \times 10^{-7}\text{ m}$$

$$t = 0.5\text{mm} = 0.5 \times 10^{-3}\text{m} = 5 \times 10^{-4}\text{ m}$$

$$\Delta y = 5\text{mm} = 5 \times 10^{-3}\text{ m}$$

The shift of interference pattern is given by

$$\Delta y = \frac{D}{2d}(\mu - 1)t$$

$$\mu - 1 = \frac{\Delta y . 2d}{D.t} = \frac{5 \times 10^{-3} \times 2 \times 10^{-3}}{0.10 \times 5 \times 10^{-4}} = 0.2$$

$$\mu = 1.2$$

Ex. 18 In Young's double-slit experiment using monochromatic light, the fringe pattern shifts by a certain distance on the screen when a mica-sheet of refractive index 1.6 and thickness 1.964 micron is introduced in the path of one of interference waves. The mica-sheet is then removed and the distance between the slits and the screen is doubled. It is found that the distance between successive maxima (or minima) now is the same as the observed fringe shift on the introduction of mica-sheet. Calculate the wavelength of the monochromatic light used in the experiment.

Sol. The shift of interference fringes is given by

$$\Delta y = \frac{D}{2d}(\mu - 1)t$$

The fringe width when distance between slits and screen is doubled is

$$\omega = \frac{(2D)\lambda}{2d}$$

According to problem, $\Delta y = \omega$,

∴ From (1) and (2), we have

$$\frac{D}{2d}(\mu - 1)t = \frac{(2D)\lambda}{2d}$$

or

$$\lambda = \frac{(\mu - 1)t}{2}$$

Here $\mu = 1.6, t = 1.964\ \mu = 1.964 \times 10^{-6}\text{m}$

$$\therefore \quad \lambda = \frac{(1.6 - 1) \times 1.964 \times 10^{-6}}{2} = 5.892 \times 10^{-7}\text{m} = 5892\ Å$$

Ex. 19 In a biprism experiment the micrometre readings for zero order and tenth order fringe are 1.25 mm. and 2.37 mm. respectively when light of $\lambda = 5.90 \times 10^{-5}$ cm is used. Now

(i) What will be the position of zero order and tenth order fringes if λ is changed to 7.50×10^{-5} cm?

(ii) If a mica sheet of thickness 0.020 mm and $\mu = 1.59$ is introduced to cover one half of the biprism, deduce the position of the zero-order fringe. (Two answers possible)

Sol. (i) The position of the zero-order fringe will remain unchanged due to change in λ, for it corresponds to zero path difference between the interfering waves of all wavelengths. Hence when λ is changed from 5.90×10^{-5}cm to 7.50×10^{-5} cm the position of zero order fringe will be 1.25 mm.

The position of other fringes will change as fringe width ω changes due to a change in λ, according to the relation

$$\omega = \frac{D\lambda}{2d} \qquad (1)$$

When, $\lambda = 5.90 \times 10^{-5}$cm, the distance between zero order and tenth order fringes

$$= 2.37 - 1.25 = 1.12 \text{ mm} = 0.112 \text{ cm}$$

i.e. $\quad 10\omega = 0.112$ cm

or $\quad \omega = \frac{0.112}{10} = 0.0112$ cm

When λ is changed to λ', let the fringe width be ω', so that we have

$$\omega' = \frac{D\lambda'}{2d}$$

$$\frac{\omega'}{\omega} = \frac{\lambda'}{\lambda}$$

$$\omega' = \frac{\lambda'}{\lambda}\omega$$

We have $\lambda' = 7.50 \times 10^{-5}$ cm, $\lambda = 5.90 \times 10^{-5}$ cm

$$\omega' = \frac{7.50 \times 10^{-8}}{5.90 \times 10^{-8}} \times 0.0112 = 0.014 \text{ cm}$$

∴ The distance between zero order and tenth order fringe is

$$10\omega' = 10 \times 0.014 \text{ cm} = 0.14\text{cm} = 1.4 \text{ mm}$$

As the position of zero order fringe is 1.25 mm. the new micrometer reading for tenth order fringe $= (1.25 + 1.4)\text{mm} = 2.65$ mm

(ii) When a mica sheet of thickness t and refractive index μ is introduced to cover one half of the biprism, the interference pattern shifts in the direction of the sheet by an amount

$$y_0 = \frac{\omega}{\lambda}(\mu - 1)t$$

Here $\mu = 1.5, t = 0.020 \text{ mm} = 0.002 \text{ cm}, \omega = 0.0112 \text{ cm}, \lambda = 5.90 \times 10^{-5}$ cm

$$y_0 = \frac{0.0112 \times (1.5 - 1) \times 0.002}{5.90 \times 10^{-5}}$$

$$= 0.19 \text{ cm} = 1.9 \text{ mm}$$

The micrometer reading for zero order fringe will now be (1.25 ± 1.9)mm i.e., either 3.15 mm or -0.65 mm. The reason for two answers is that the central fringe may shift on either side depending upon the direction of introduction of the plate.

Based on Thin film interference

Ex. 20 (i) A thin layer of colourless clean oil is spread over the water in a container. The white light incident on the surface appears green after reflection. Why?

(ii) If the light of wavelength 6400 Å be absent in the reflected light, what should be the minimum thickness of oil layer?

Sol. (i) The conditions of maxima and minima in reflected light are

$$2\mu t \cos r = (2n - 1)\frac{\lambda}{2} (\text{maxima})$$

$$2\mu t \cos r = n\lambda (\text{minima})$$

The reflected light appears green if the red colour is absent from reflected light; then (white-red) appears green. (Refractive index of oil = 1.4).

(ii)The thickness t satisfying the condition of minima in reflected light is given by

$$2\mu t \cos r = n\lambda$$

$$t = \frac{n\lambda}{2\mu \cos r}$$

For minimum thickness, $n_{min} = 1, (\cos r)_{max} = 1$

$$t_{min} = \frac{\lambda}{2\mu} = \frac{6000Å}{2 \times 1.4} = 2143Å$$

Ex.21 A glass slab ($\mu = 1.50$) is to be coated with a film of some transport material ($\mu = 1.25$). What should be the thickness of the film yellow light $(\lambda = 6000Å)$ incident normally on the film is not reflected?

Sol. In this case of interference in thin films, the situation is somewhat different. The reflections at both the upper and lower surface of the material ($\mu = 1.25$) film take place under similar conditions i.e., when light is going from a rarer to a denser medium. Thus, there is a phase change of π at both reflections which means no phase difference due to reflection between the two interfering beams.

The path difference between the two interfering beams is $2\mu t$ for normal incidence, where t is the thickness and μ the refractive index of the film.

The two beams will destroy each other if the path difference is an odd multiple of $\frac{\lambda}{2}$, i.e., when

$$2\mu t = (2n-1)\frac{\lambda}{2} \qquad n = 1,2,3,.$$

This is the condition of minima.

Here $\mu = 1.5$ and $\lambda = 6000Å$

$$2 \times 1.25 \times t = (2n-1) \times \frac{6000}{2}Å$$

Hence the required thickness is given by

$$t = (2n-1)\frac{6000}{2 \times 2 \times 1.25}Å$$

$$= (2n-1)1200Å; \qquad \text{where } n = 1,2,3,\ldots$$

Ex.22 Calculate the minimum thickness of MgF_2 film ($\mu = 1.38$) coating to reduce reflection from the glass surface for light of wavelength 5500Å using interference techniques.

Sol. Minimum thickness, $t_{min} = \frac{\lambda}{4\mu_f} = \frac{5500}{4 \times 1.38} = 1000Å$

Ex. 23 Light of wavelength 5893 Å is reflected at nearly normal incidence from a soap film of refractive index 1.42. What is the least thickness of the film that will appear (i) black, (ii) bright?

Sol. (i) The condition for the blackness of the film in reflected light is

$$2\mu t \cos t = n\lambda$$

For normal incidence, $r = 0$ and $\cos r = 1$

$$2\mu t = n\lambda$$

$$t = \frac{n\lambda}{2\mu}$$

For least thickness of the film, $n = 1$

$\therefore$ Least thickness of the film, $t = \frac{\lambda}{2\mu}$

$$= \frac{5893 \times 10^{-8}}{2 \times 1.42} \text{ cm}$$

$$= 2075 \times 10^{-8} \text{ cm} = 2075Å$$

(ii) The condition for brightness of the film in reflected light is

$$2\mu t \cos r = (2n-1)\lambda/2$$

For normal incidence r = 0; so cos r = 1

$$2\mu t = (2n-1)\lambda/2$$

$$t = \frac{(2n-1)\lambda}{2\mu \quad 2} = \frac{(2n-1)\lambda}{4\mu}$$

or for least thickness of the film, n = 1

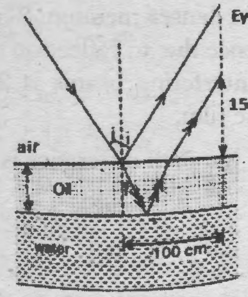
∴ Least thickness of the film, $t = \frac{\lambda}{4\mu}$

$$= \frac{5893 \times 10^{-8}}{4 \times 1.42} cm$$

$$= 1037.5 \times 10^{-8} \text{ cm} = 1037.5Å$$

Ex.24 A man whose eyes are 150 cm above the oil film on water surface observes greenish colour at a distance of 100 cm from his feet. Calculate the probable thickness of the film.

$$\lambda_{green} = 5000Å, \quad \mu_{oil} = 1.4, \quad \mu_{water} = 1.33$$



Sol. The condition for maxima in the reflected light is given by.

$$2\mu t \cos r = (2n-1)\frac{\lambda}{2},$$

$$t = \frac{(2n-1)\lambda}{4\mu \cos r} \qquad (1)$$

Here λ=5000Å= 5×10^{-5} cm, $\mu_{oil} = 1.4$ and from Fig.

$$\tan i = \frac{100}{150} = \frac{2}{3}$$

∴ $$\sin i = \frac{2}{\sqrt{13}}$$

But $$\mu = \frac{\sin i}{\sin r}$$

∴ $$\sin r = \frac{\sin i}{\mu} = \frac{2/\sqrt{13}}{1.4} = 0.3962$$

∴ $$\cos r = \sqrt{1-\sin^2 r} = \sqrt{1-(0.3962)^2}$$

$$= 0.9181$$

Substituting the values in equation (1), the probable thickness of the film is

$$t = \frac{(2n-1) \times 5 \times 10^{-5}}{4 \times 1.4 \times 0.9181} = 9.725 \times 10^{-6}(2n-1) \text{ cm} \qquad n = 1,2,3\ldots\ldots$$

Ex.25 A parallel beam of sodium light $(\lambda = 5890Å)$ strikes a film of oil floating on water. When viewed at an angle of 30° from the normal, 8th dark band is seen. Determine the thickness of the film. (Refractive index of oil = 1.46)

Sol. According to the condition of dark bands for a film seen in reflected light

$$2\mu t \cos r = n\lambda$$

where r is the angle of refraction in the film.

$\therefore \quad t = \dfrac{n\lambda}{2\mu\cos r'}$

We have $\quad \mu = \dfrac{\sin i}{\sin r'}$

$\therefore \quad \sin r = \dfrac{\sin i}{\mu}$

So that $\quad \cos r = \sqrt{1-\sin^2 r} = \sqrt{\left(1-\dfrac{\sin^2 i}{\mu^2}\right)}$

Here $\mu = 1.46, i = 30°$

$$\cos r = \sqrt{1-\frac{\sin^2 30°}{(1.46)^2}}$$

$$= \sqrt{1-\frac{(1/2)^2}{(1.46)^2}}$$

$$= \sqrt{1-\frac{1}{4\times(1.46)^2}} = 0.94$$

$\therefore$ From equation (1), $\quad t = \dfrac{8\times 5890\times 10^{-10}}{2\times 1.46\times 0.94}$ (since $n = 8, \lambda = 5890\times 10^{-10}$m)

$= 1.7\times 10^{-6}$ m

Ex.26 White light falls perpendicularly upon a film of soapy water whose thickness is 5×10^{-5} cm and whose index of refraction is 1.33. Which wavelength in the visible region will be reflected most strongly?

Sol. The condition of maximum intensity in the light reflected from a thin film is

$$2\mu t\cos r = (2n-1)\frac{\lambda}{2} \qquad (1)$$

Here $\mu = 1.33, t = 5\times 10^{-5}$ cm

$r = 0°$ (Since light on the film falls normally) so that $\cos r = 1$

$\therefore$ Equation (1) yields

$$2\times 1.33\times 5\times 10^{-5}\times 1 = (2n-1)\frac{\lambda}{2}$$

or

$$\lambda = \frac{2\times 2\times 1.33\times 5\times 10^{-5}}{2n-1}$$

$$= \frac{2.66\times 10^{-4}}{2n-1} = \frac{26600}{2n-1}\times 10^{-8}\text{ cm}$$

when $n = 1$, $\quad \lambda = 26600\times 10^{-8}$ cm $= 26600$Å

when $n = 2$, $\quad \lambda = \dfrac{26600\times 10^{-8}}{3}$ cm $= 8866.66\times 10^{-8}$cm $= 8866.66$Å

when $n = 3$, $\quad \lambda = \dfrac{26600\times 10^{-8}}{5}$ cm $= 5320$Å

when $n = 4$, $\quad \lambda = \dfrac{26600\times 10^{-8}}{7}$ cm $= 3800$Å and so on.

These wavelengths are reflected most strongly. Out of these only the wavelength 5320 Å. lies in the visible region (4000 Å – 7500 Å). Therefore, in the visible region the wavelength 5320 Å is reflected most strongly.

Ex.27 White light is incident on a soap film at an angle $\sin^{-1}\frac{4}{3}$ and the reflected light is observed with a spectroscope. It is found that two consecutive dark bands correspond to wavelengths 6.1×10^{-5} and 6.0×10^{-5} cm . If the refractive index of the film be $\frac{1}{5}$, calculate its thickness.

Sol. The condition for dark bands in the reflected light is

$$2\mu t\cos r = n\lambda$$

Two consecutive dark bands differ in order by 1. If n and (n+1) are the orders of the consecutive dark bands for wavelengths λ_1 and λ_2 respectively, we must have

$$2\mu t\cos r = n\lambda_1 \qquad (1)$$

and

$$2\mu t\cos r = (n+1)\lambda_2 \qquad (2)$$

From (1) and (2)

$$n\lambda_1 = (n+1)\lambda_2$$

or

$$n\lambda_1 = n\lambda_2 + \lambda_2$$

or

$$n(\lambda_1 - \lambda_2) = \lambda_2$$

$\therefore$

$$n = \frac{\lambda_2}{\lambda_1-\lambda_2}$$

Substituting this value of n in equation (1), we get

$$2\mu t\cos r = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}$$

$$t = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}\cdot\frac{1}{2\mu\cos r} \qquad (3)$$

But

$$\cos r = \sqrt{(1-\sin^2 r)}$$

$$= \sqrt{\left\{1-\left(\frac{\sin i}{\mu}\right)^2\right\}} \qquad \left(\text{since } \mu = \frac{\sin i}{\sin r}\right)$$

$$= \frac{\sqrt{\mu^2-\sin^2 i}}{\mu}$$

$\therefore$ Substituting this value of cos r in (3), we get

$$t = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}\cdot\frac{1}{2\sqrt{(\mu^2-\sin^2 i)}}$$

Here $\lambda_1 = 6.1\times10^{-5}$ cm, $\lambda_2 = 6.0\times10^{-5}$cm, $\mu = \frac{4}{3}$ and $\sin i = \frac{4}{5}$

$$t = \frac{6.1\times10^{-5}\times6.0\times10^{-3}}{(6.1\times10^{-5}\times6.0\times10^{-5})}\times\frac{1}{2\sqrt{\left[\left(\frac{4}{3}\right)^2-\left(\frac{4}{5}\right)^2\right]}}$$

$$= \frac{6.1\times10^{-5}\times6.0\times10^{-3}}{0.1\times10^{-5}}\times\frac{1}{2\sqrt{\left(\frac{16}{9}-\frac{16}{25}\right)}}$$

$$= \frac{6.1\times6.0\times10^{-5}}{0.1\times2\times\left(\frac{16}{15}\right)}$$

$$= \frac{6.1\times6.0\times10^{-5}\times15}{0.1\times2\times16}$$

$$= 0.0017 \text{ cm}$$

Ex. 28 A parallel beam of white light is allowed to fall axially on the slit of the collimator of a spectrometer and a thin air film of uniform thickness enclosed between two glass-plates is placed in front of the slit with the surface of the air film perpendicular to the path of light. On viewing through the telescope 250 fringes are observed in the spectral region between wavelength 4000 Å and 6500 Å as measured in air. Calculate the thickness of the air film.

Sol. The condition of maxima when seen by the transmitted light is

$$2\mu t\cos r = n\lambda$$

In this case the air film is seen by light transmitted normally through it, i.e., here r = 0

$$2\mu t = n\lambda \qquad (1)$$

Let n be the order of bright fringe corresponding to light of wavelength 6500. Å. Then the order of bright fringe corresponding to light of wavelength 4000 Å will be (n+250); because according to this problem 250 fringes are observed in the spectral region between these two wavelengths.

∴ According to equation (1), we have

$$2\mu t = n \times 6500 \times 10^{-8} \quad (2)$$

and

$$2\mu t = (n + 250) \times 4000 \times 10^{-8} \quad (3)$$

Comparing equations (2) and (3), we have

$$n \times 6500 \times 10^{-8} = (n + 250) \times 4000 \times 10^{-8}$$

or

$$65n = 40n + 40 \times 250$$

or

$$25n = 40 \times 250$$

or

$$n = \frac{40\times250}{25} = 400$$

From equation (2), we have

$$2\mu t = n \times 6500 \times 10^{-8}$$

Here μ=1(for air) and n=400

$$2t = 400 \times 6500 \times 10^{-8}$$

or

$$t = \frac{400 \times 6500 \times 10^{-8}}{2} = 0.013 \text{ cm}$$

Ex.29 White light is reflected from an oil film of thickness 0.01 mm and refractive index 1.4 at an angle of 45° to the vertical. If the reflected light falls on the slit of a spectrometer, calculate the number of dark bands seen between wavelengths 4000 and 5000 Å.

Sol. The condition of dark bands in reflected light is

$$2\mu t \cos r = n\lambda$$

For wavelength $\lambda_1 (= 4000\text{Å})$, we have

$$2\mu t \cos r = n_1\lambda_1$$

$$n_1 = \frac{2\mu t \cos r}{\lambda_1}$$

For wavelength $\lambda_2 (= 5000\text{Å})$,

$$2\mu t \cos r = n_2\lambda_2$$

$$n_2 = \frac{2\mu t \cos r}{\lambda_2}$$

Here $\mu = 1.4, t = 0.01 \text{ mm} = 0.001 \text{ cm}$, angle of reflection = angle of incidence $i = 45°$

We have

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

i.e.

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

$$\cos r = \sqrt{(1 - \sin^2 r)}$$

$$= \sqrt{1 - \frac{\sin^2 i}{\mu^2}}$$

$$= \sqrt{1 - \frac{\sin^2 45°}{\mu^2}}$$

$$= \sqrt{1 - \frac{1}{2\mu^2}}$$

$$= \sqrt{1 - \frac{1}{2 \times (1.4)^2}} = 0.86$$

∴

$$n_1 = \frac{2\mu t \cos r}{\lambda_1} = \frac{2 \times 1.4 \times 0.001 \times 0.86}{4000 \times 10^{-8}} \approx 60$$

and $$n_2 = \frac{2\mu t \cos r}{\lambda_2} = \frac{2 \times 1.4 \times 0\ .001 \times 0.86}{5000 \times 10^{-8}} \approx 48$$

The number of dark bands seen within wavelengths 4000 Å and 5000 Å is

$$n_1 - n_2 = 60 - 48 = 12.$$

Based on Wedge Shaped Films

Ex. 30 A square piece of cellophane film with index of refraction 1.5 has a wedge-shaped section so that its thickness of two opposite sides is t_1 and t_2. If with a light of $\lambda = 6000$ Å the number of fringes appearing on the film is 10, calculate the difference $t_1 - t_2$.

Sol. Let the order of the fringe appearing at one end of the film be n. Then the order of the fringe appearing at another end will be (n+10). Therefore, we have for n^{th} and $(n+10)^{th}$ dark fringes,

$$2\mu t_1 \cos(r + \theta) = n\lambda \qquad (1)$$

and $$2\mu t_1 \cos(r + \theta) = (n + 10)\lambda \qquad (2)$$

where θ is the angle of wedge and r is the angle of refraction inside the film.

Subtracting (1) from (2), we get

$$2\mu(t_2 - t_1)\cos(r + \theta) = 10\lambda \qquad (3)$$

If the fringe is seen normally and the angle of wedge is very small, then r = 0, so that

$$\cos(r + \theta) = \cos\theta = 1$$

Then equation (3) gives, $2\mu(t_2 - t_1) = 10\lambda$

$$t_2 - t_1 = \frac{10\lambda}{2\mu}$$

$$= \frac{5\lambda}{\mu} = \frac{5 \times 6000 \times 10^{-10}}{1.5} \quad (\text{Since } \mu = 1.5 \text{ and } \lambda = 6000\ \text{Å given})$$

$$= 2 \times 10^{-6}\text{m} = 2\ \mu\text{m}$$

Ex.31 Using sodium light (λ=5893 Å) interference fringes are formed from a thin air wedge. When viewed normally 10 fringes are observed in a distance of 1 cm. Calculate the angle of the wedge.

Sol. In the case of wedge-shaped film of wedge angle θ, refractive index μ, the fringe which ω for normal incidence is

$$\omega = \frac{\lambda}{2\mu\theta}, \ \lambda \text{ being the wavelength of light used}$$

$$\therefore \qquad \theta = \frac{\lambda}{2\mu\omega}$$

Here $\lambda = 5893\text{Å} = 5893 \times 10^{-8}$ cm, $\mu = 1$ (for air)

and fringe width, $$\omega = \frac{\text{distance given}}{\text{number of fringes in given distance}} = \frac{1}{10} = 0.1 \text{ cm}$$

$$\theta = \frac{5893 \times 10^{-8}}{2 \times 1 \times 0.1}$$

$$= 2.946 \times 10^{-4} \text{ radians}$$

$$= 2.946 \times 10^{-4} \times \frac{180}{\pi} \text{ degree}$$

$$= \frac{2.946 \times 10^{-4} \times 180}{3.14} \times 60 \times 60 \text{ sec}$$

$$= 61 \text{ sec} = 1 \text{ min } 1 \text{ sec}$$

Ex. 32 Two optically plane glass strips of length 10 cm are placed one over the other. A thin foil of thickness 0.010 mm is introduced between the plates at one end to form an air film. If the light used has wavelength 5900 Å, find the separation between consecutive bright fringes.

Sol. The angle of wedge $$\theta = t/x \qquad (1)$$

where x is the length of the film and t is the thickness of the foil.

The fringe width $\omega = \lambda/(2\mu\theta)$

Substituting value of θ from equation (1), we get

$$\omega = \frac{\lambda x}{2\mu t}$$

Here $\lambda = 5900\text{Å} = 5900 \times 10^{-8}$ cm , $x = 10$ cm, $t = 0.010$ mm $= 0.001$ cm and $\mu = 1$ (for air)

$$\omega = \frac{5900 \times 10^{-8} \times 10}{2 \times 1 \times 0.001} \text{ cm.} = 0.295 \text{ cm}$$

Ex.33 Two rectangular pieces of a plane glass are laid one upon the other and thin wire is placed between them so that a thin wedge of air is formed between them. The plates are illuminated with sodium light (λ = 5893 Å) at normal incidence. Bright and dark bands are formed, there being 10 of each per centimetre length of the wedge measured normal to the edge in contact. Find the angle of the wedge.

Sol. If the thickness of the film at a distance x from the edge is t then

angle of wedge $\theta = t/x$ (1)

For n^{th} dark fringe, we have $2\mu t \cos(r + \theta) = n\lambda$. (2)

and for $(n+m)^{th}$ dark fringe, we have

$$2\mu t' \cos(r + \theta) = (n + m)\lambda \quad (3)$$

Subtracting (2) from (3), we get

$$2\mu(t' - t)\cos(r + \theta) = m\lambda$$

or $2\mu(x'\theta - x\theta)\cos(r + \theta) = m\lambda$ [using equation (1)]

or $2\mu(x' - x)\theta\cos(r + \theta) = m\lambda$ (4)

For normal incidence $r = 0$

$$\cos(r + \theta) = \cos\theta = 1 \text{ (For small angle of wedge)}$$

∴ Equation (4) gives

$$2\mu(x' - x)\theta = m\lambda$$

or $\theta = \frac{m\lambda}{2\mu(x'-x)}$ (5)

Here $\lambda = 5893\text{Å} = 5893 \times 10^{-8}$ cm , $\mu = 1$ (for air), $m = 10$ and $x' - x = 1$ cm

Thus from (5),

$$\theta = \frac{10 \times 5893 \times 10^{-8}}{2 \times 1 \times 1} = 2.9465 \times 10^{-4} \text{ radians}$$

Ex.34 A soap film viewed in reflected light in nearly normal direction shows red colour at one end. Then successive colours yellow, green, blue and again red is seen. If the distance between consecutive red colours is 6 mm., what information do we get about the film from these observations? ($\lambda_{red} = 6.5 \times 10^{-5}$cm , $\lambda_{blue} = 4.3 \times 10^{-5}$cm , n and $\mu_{water} = 1.33$)

Sol. If θ is the angle of wedge, then the fringe width ω for wavelength λ is given by

$$\omega = \frac{\lambda}{2\mu\theta}$$

$$\therefore \quad \theta = \frac{\lambda}{2\mu\omega}$$

where μ is the refractive index of the film. The fringe width, in this case, is given for red light, i.e. $\omega_{red} = 6\text{mm} = 0.6$ cm

$$\theta = \frac{\lambda_{red}}{2\mu\omega_{red}}$$

$$= \frac{6.5 \times 10^{-5}}{2 \times 1.33 \times 0.6} \quad \text{(since } \mu = 1.33)$$

$$= 4.07 \times 10^{-5} \text{ radians}$$

For normal incidence the thickness of the film t is given by

$$2\mu t = (2n-1)\frac{\lambda}{2}$$

or
$$t = \frac{(2n-1)\lambda}{4\mu} \qquad (1)$$

As the value in the visible region increases from violet to red, being maximum for red colour and minimum for violet colour and in this case, we are given that first we obtain red colour, then yellow, green, blue etc. this means in the film the condition of maxima is first satisfied by red light, then by yellow, green, blue etc. This is only possible if the film is of decreasing thickness. Thus, from the above given information:

(i)The film is of decreasing thickness and ii) the angle of wedge is 4.07×10^{-5} radians.

Based On Newton's Rings

Ex.35 (a) Show that with a spherical convex surface and a plane surface, the squares of the diameters of successive rings have equal difference. What will be the rings pattern if the concave surface of the plano-concave lens is placed towards the plane surface?

(b) What will happen if in Newton's ring arrangement we use mirror instead of glass plate below the lens?

(c) What will happen if we use a lens of small radius of curvature?

Sol. (a) The diameter of n^{th} dark ring is given by

$$D_n^2 = 4n\lambda R \qquad (1)$$

where λ is wavelength of light used and R is the radius of the convex surface of the plano convex lens.

The diameter of $(n+1)^{th}$ dark ring is given by

$$D_{n+1}^2 = 4(n+1)\lambda R \qquad (2)$$

Subtracting (1) from (2), we get

$$D_{n+1}^2 - D_n^2 = 4\lambda R \qquad (3)$$

From equation (3) it is clear that the difference between the squares of the diameters of successive dark rings is independent of the order n. This means that with a spherical convex surface and a plane surface, the squares of diameters of successive dark rings have equal difference. This result may also be proved for bright rings.

Ex.36 Newton's rings are formed by reflection in the air film between a plane surface and spherical surface of radius 50 cm, and it is noticed that the centre of the ring system is bright. If the diameter of the third bright ring is 0.181 cm. and the diameter of the 23rd bright ring is 0.501 cm. What is the wavelength of light used? What conclusion would be drawn from the fact that the centre of the ring system is bright?

Sol. The wavelength of light,
$$\lambda = \frac{D_{n+p}^2 - D_n^2}{4pR} \qquad (1)$$

Given diameter of 3rd ring, $D_n = D_3 = 0.181$ cm

and diameter of 23rd ring $D_{n+p} = D_{23} = 0.501$ cm

$\therefore$ $p = (n+p) - n = (23-3) = 20$

and radius of curvature of the lens, R = 50 cm.

Substituting these values in equation (1), we get

$$\lambda = \frac{(0.501)^2 - (0.181)^2}{4 \times 20 \times 50} \text{cm} = 5456 \times 10^{-8} \text{cm} = 5456 \text{Å}$$

As the centre of the ring system in the reflected light is bright, this means that the lens is not perfectly in contact with plate, due to which a path difference of the order $\frac{\lambda}{2}$ exists.

For normal incidence the condition for brightness in reflected light for air film ($\mu = 1$)is,

$$2t = (2n - 1)\lambda/2.$$

The central spot corresponds to $n = 1$,

$\therefore$ $$2t = \lambda/2$$

or $$t = \lambda/4$$

This means that an air film of thickness $\lambda/4$ exists between the centre of the lens and the plate.

Ex. 37 In a Newton's ring experiment the diameter of 4th and 12th dark rings are 0.400 cm. and 0.700 cm respectively. Deduce the diameter of 20th dark ring.

Sol. If D_{n+p} and D_n are the diameters of $(n + p)^{th}$ and n^{th} rings respectively, then we have

$$D_{n+p}^2 - D_n^2 = 4p\lambda R$$

First putting $n = 4$ and $n + p = 12$, so that

$$D_n = D_4 = 0.4 \text{ cm}$$

$$D_{n+p} = D_{12} = 0.7 \text{ cm}$$

Now equation (1)in the above example gives

$\therefore$ $$D_{12}^2 - D_4^2 = 4 \times (12 - 4) \times \lambda R = 4 \times 8 \times \lambda R \quad (2)$$

Again putting $n = 4$ and $n + p = 20$, we get

$$D_{20}^2 - D_4^2 = 4 \times 16 \times \lambda R \quad (3)$$

Dividing (2) by (3), we get

$$\frac{D_{12}^2 - D_4^2}{D_{20}^2 - D_4^2} = \frac{4 \times 8}{4 \times 16} = \frac{1}{2} \quad (4)$$

Given $D_4 = 0.400\text{cm}$, $D_{12} = 0.700$ cm, therefore equation (4) gives

$$\frac{(0.700)^2 - (0.400)^2}{D_{20}^2 - (0.400)^2} = \frac{1}{2}$$

or $$\frac{0.49 - 0.16}{D_{20}^2 - 0.16} = \frac{1}{2}$$

or $$\frac{0.33}{D_{20}^2 - 0.16} = \frac{1}{2}$$

or $$D_{20}^2 - 0.16 = 0.66$$

or $$D_{20}^2 = 0.66 + 0.16 = 0.82$$

Diameter of 20th dark ring, $D_{20} = \sqrt{0.82} = 0.905$ cm

Ex.38 A Newton's rings arrangement is used with source emitting two wavelength $\lambda_1 = 6000\text{Å}$ and $\lambda_2 = 4500\text{Å}$ and it is found that n^{th} dark ring due to λ_1 coincides with $(n + 1)^{th}$ dark ring due to λ_2. If the radius of curvature of the curved surface is 90 cm. Find the diameter of n^{th} and $(n + 5)^{th}$ dark ring for λ_2.

Sol. Let D be the diameter of nth dark ring for λ_1. According to given problem D is also the diameter of $(n + 1)^{th}$ dark ring for λ_2.

We know $$D_n^2 = \frac{4n\lambda R}{\mu} \quad (1)$$

$\therefore$ For λ_1, $$D^2 = D_n^2 = \frac{4n\lambda_1 R}{\mu} \quad (2)$$

or λ_2 $$D^2 = D_{n+1}^2 = \frac{4(n+1)\lambda_2 R}{\mu} \quad (3)$$

Comparing (2) and (3), we get

$$n\lambda_1 = (n + 1)\lambda_2 \quad (4)$$

or $$\frac{n+1}{n} p = \frac{\lambda_1}{\lambda_2}$$

or
$$1+\frac{1}{n}=\frac{\lambda_1}{\lambda_2}$$
or
$$\frac{1}{n}=\frac{\lambda_1}{\lambda_2}-1=\frac{\lambda_1-\lambda_2}{\lambda_2}$$
$\therefore$
$$n=\frac{\lambda_2}{\lambda_1-\lambda_2}$$
Given $\lambda_2 = 4500Å = 4500 \times 10^{-10}\ m = 4.5 \times 10^{-7}\ m$

$\lambda_1 = 6000Å = 6000 \times 10^{-10} m = 6 \times 10^{-7}\ m$

$\therefore$
$$n=\frac{4.5\times10^{-7}}{6\times10^{-7}-4.5\times10^{-7}}$$
$$=\frac{4.5\times10^{-7}}{1.5\times10^{-7}}=3.$$
$\therefore$ Diameter of n^{th} dark ring D_n for λ_1 is given by
$$D_n^2 = 4n\lambda_1 R = 4\times3\times6\times10^{-7}\times0.90$$
$$D_n=\sqrt{(4\times3\times6\times10^{-7}\times0.90)}=2.545\times10^{-3}\ m=2.545\ mm$$
$$D_{n+3}=\sqrt{4(n+3)\lambda_1R}$$
$$=\sqrt{(4\times6\times6\times10^{-7}\times0.90)}=3.600\times10^{-3}\ m=3.600\ mm$$

Ex.39 Newton's rings are formed in reflected light of wavelength 6000Å with a liquid between the plane and curved surfaces. If the diameter of the sixth bright ring be 3.1 mm. and the radius of curvature of the curved be 1 m. Calculate the refractive index of the liquid.

Sol. The diameter of n^{th} bright ring is given by
$$D_n^2=\frac{2(2n-1)\lambda R}{\mu}$$
or
$$\mu=\frac{2(2n-1)\lambda R}{D_n^2}$$
Given $n = 6, \lambda = 6000Å = 6000 \times 10^{-10} m, R = 1m$, and $D_n = 3.1mm = 3.1 \times 10^{-3}m$
$$\mu=\frac{2\times(12-1)\times6000\times10^{-10}\times1}{(3.1\times10^{-3})^2}$$
$$=\frac{2\times11\times6000\times10^{-10}\times1}{9.61\times10^{-6}}=1.374$$

Ex. 40 If the diameter of n^{th} dark ring in an arrangement giving Newton's rings changes from 0.30 cm. to 0.25 cm. as a liquid is introduced between the lens and the plate. Calculate the value of the refractive index of the liquid.

Sol. The diameter of nth dark ring in the presence of the liquid (below the convex surface) of refractive index μ is given by
$$(D_n^2)_{liquid}=\frac{4n\lambda R}{\mu} \qquad (1)$$
In the absence of the liquid (i.e. for air film $\mu = 1$)
$$(D_n^2)_{air}=4nR \qquad (2)$$
From (1) and (2) $\therefore$
$$\mu=\frac{(D_n^2)_{air}}{(D_n^2)_{liquid}}$$
$$=\left\{\frac{(D_n^2)_{air}}{(D_n^2)_{liquid}}\right\}^2$$
$$=\left[\frac{0.30}{0.25}\right]^2=1.44$$
$$D_{n+3}=\sqrt{4(n+3)\lambda_1R}$$

$$= \sqrt{4 \times 6 \times 6 \times 10^{-7} \times 0.90}$$
$$= 3.60 \times 10^{-3}\text{m} = 3.60 \text{ mm}$$

Ex.41 In a Newton's ring arrangement with air film observed with light of wavelength 6×10^{-5}cm , the difference of squares of diameters of successive rings are 0.125 cm². What will happen to this quantity if:

(i) wavelength of light is changed to 4.5×10^{-5} cm

(ii) a liquid of refractive index 1.33 is introduced between the lens and the plate;

(iii) the radius of curvature of the convex surface of the plano convex lens is doubled.

Sol. If D_n and D_{n+p} are the diameter of nth and (n+p)th rings, we have

$$D_{n+p}^2 - D_n^2 = \frac{4p\lambda R}{\mu}$$

Here $p = 1,$
$$D_{n+1}^2 - D_p^2 = \frac{4\lambda R}{\mu} \qquad (1)$$

(i) When wavelength of light is changed from λ to λ_1, we have

$$(D_{n+1}^2 - D_n^2)_{\lambda_1} = \frac{4\lambda_1 R}{\mu} \qquad (2)$$

Dividing (2) by (1), we get

$$\frac{(D_{n+1}^2 - D_n^2)_{\lambda_1}}{D_{n+1}^2 - D_n^2} = \frac{\lambda_1}{\lambda}$$

$$\therefore \qquad (D_{n+1}^2 - D_n^2)_{\lambda_1} = \frac{\lambda_1}{\lambda}(D_{n+1}^2 - D_n^2)$$

Here $\lambda_1 = 4.5 \times 10^{-5}$cm , $\lambda = 6 \times 10^{-5}$cm , $(D_{n+1}^2 - D_n^2) = 1.125 \text{ cm}^2$

$$\therefore \qquad (D_{n+1}^2 - D_n^2)_{\lambda_1} = \frac{4.5 \times 10^{-5}}{6 \times 10^{-5}} \times 0.125 = 0.094\text{cm}^2.$$

(ii) When a liquid of refractive index μ_1 is introduced between the lens and the plate, we have

$$(D_{n+1}^2 - D_n^2)_{\mu_1} = \frac{4\lambda R}{\mu_1} \qquad (3)$$

Dividing (3) by (1), we get

$$\frac{(D_{n+1}^2 - D_n^2)\mu_1}{D_{n+1}^2 - D_n^2} = \frac{\mu}{\mu_1}$$

or
$$(D_{n+1}^2 - D_n^2)_{\mu_1} = \frac{\mu}{\mu_1}(D_{n+1}^2 - D_n^2)$$

Here, $\mu = 1, \mu_1 = 1.33$ and $D_{n+1}^2 - D_n^2 = 0.125 \text{ cm}^2$

$$\therefore \qquad (D_{n+1}^2 - D_n^2)_{\mu} = \frac{1}{1.33} \times 0.125 = 0.094\text{cm}^2$$

(iii) When the radius of the convex surface is changed to R_1, we have

$$(D_{n+1}^2 - D_n^2)_{R_1} = \frac{4\lambda R_1}{\mu} \qquad (4)$$

Dividing equation (4) by (1), we get

$$\frac{(D_{n+1}^2 - D_n^2)_{R_1}}{D_{n+1}^2 - D_n^2} = \frac{R_1}{R}$$

$$(D_{n+1}^2 - D_n^2)_{R_1} = \frac{R_1}{R} \times (D_{n+1}^2 - D_n^2)$$

$$(D_{n+1}^2 - D_n^2)_{2R} = \frac{2R}{R} \times 0.125\text{cm}^2 = 2 \times 0.125 = 0.250 \text{ cm}^2$$

Ex.42 The convex surface of radius 300 cm. of a plano-convex lens rests on the concave spherical surface of radius 400 cm. If Newton's rings are viewed with reflected light of wavelength 6×10^{-5}cm.

Calculate (i) the diameter of the 12th dark ring and 13th bright ring. (ii) the difference of squares of the diameter of successive rings.

Sol. (i) The diameter of nth dark ring D_n is given as:

$$D_n^2 = \frac{4n\lambda}{\left[\frac{1}{R_1} - \frac{1}{R_2}\right]} \quad (1)$$

Here $n = 12, \lambda = 6 \times 10^{-5}$cm, $R_1 = 300$ cm and $R_2 = 400$ cm

$$D_{12}^2 = \frac{4 \times 12 \times 6 \times 10^{-5}}{\left[\frac{1}{300} - \frac{1}{400}\right]}$$

$$= \frac{4 \times 12 \times 6 \times 10^{-5}}{(1/1200)} \text{cm}^2$$

$$= 4 \times 12 \times 6 \times 10^{-5} \times 1200 \text{ cm}^2$$

∴ diameter of 12th dark ring

$$D_{12} = \sqrt{(4 \times 12 \times 6 \times 10^{-5} \times 1200)} = 1.857 \text{ cm}$$

The diameter of nth bright ring D_n is given by

$$D_n^2 = \frac{2(2n-1)\lambda}{\left[\frac{1}{R_1} - \frac{1}{R_2}\right]}$$

Here $n = 13, \lambda = 6 \times 10^{-5}$ cm, $R_1 = 300$ cm and $R_2 = 400$ cm

$$D_{13}^2 = \frac{2 \times 25 \times 6 \times 10^{-5}}{\left[\frac{1}{300} - \frac{1}{400}\right]}$$

$$= 2 \times 25 \times 6 \times 10^{-5} \times 1200 \text{cm}^2$$

∴ the diameter of 13th bright ring,

$$D_{13} = \sqrt{(2 \times 25 \times 6 \times 1200 \times 10^{-5})} \text{ cm}$$

If D_n and D_{n+1} are the diameters of nth and (n + 1)th dark rings, we have

$$D_n^2 = \frac{4n\lambda}{\left[\frac{1}{R_1} - \frac{1}{R_2}\right]}$$

and

$$D_{n+1}^2 = \frac{4(n+1)\lambda}{\left[\frac{1}{R_1} - \frac{1}{R_2}\right]}$$

∴

$$D_{n+1}^2 - D_n^2 = \frac{4\lambda}{\left[\frac{1}{R_1} - \frac{1}{R_2}\right]}$$

$$= \frac{4 \times 6 \times 10^{-5}}{\left[\frac{1}{300} - \frac{1}{400}\right]}$$

$$= 4 \times 6 \times 10^{-5} \times 1200 = 0.288 \text{cm}^2$$

Ex.43 Newton's rings by reflection are formed between two biconvex lenses having equal radii of curvature being 100 cm each. Calculate the distance between the 5th and 15th dark ring using monochromatic light of wavelength 5400 Å.

Sol. When the air film is formed between the biconvex lens, we have

$$r_n^2 = \frac{n\lambda}{\left[\frac{1}{R_1} + \frac{1}{R_2}\right]}$$

where r_n is the radius of nth dark ring.

For 5th dark ring, n = 5 and for 15th dark ring, n = 15

$$r_5^2 = \frac{5\lambda}{\left[\frac{1}{R_1} + \frac{1}{R_2}\right]}$$

and

$$r_{15}^2 = \frac{15\lambda}{\left[\frac{1}{R_1} + \frac{1}{R_2}\right]}$$

Given $\lambda = 5400 \times 10^{-8}$cm , $R_1 = R_2 = 100$ cm

$\therefore$ $$r_5^2 = \frac{5 \times 5400 \times 10^{-8}}{\left(\frac{1}{100} + \frac{1}{100}\right)} = \frac{5 \times 5400 \times 10^{-8} \times 100}{2}$$

This gives $$r_5 = \sqrt{\left(\frac{5 \times 5400 \times 10^{-8} \times 100}{2}\right)} = 0.1161 \text{ cm}$$

Now $$r_{15}^2 = \frac{15 \times 5400 \times 10^{-8}}{\left(\frac{1}{100} + \frac{1}{100}\right)} = \frac{15 \times 5400 \times 10^{-8} \times 100}{2}$$

$$r_5 = \sqrt{\left(\frac{5 \times 5400 \times 10^{-8} \times 100}{2}\right)} = 0.2 \text{ cm}$$

$\therefore$ The distance between the 5th and 15th dark rings

$$= r_{15} - r_5$$
$$= 0.2000 - 0.1161 = 0.0839 \text{ cm}$$

Based On Michelson Interferometer

Ex.44 When the movable mirror in a Michelson Interferometer is moved through 0.20mm 800 rings cross the field of view. What is the wavelength of light used?

Sol. $\lambda = ?$, $n = 800$, $(x_2 - x_1) = 0.20$mm

$$2(x_2 - x_1) = n\lambda$$

Or $$\lambda = \frac{2(x_2 - x_1)}{n} = \frac{2 \times 0.2 \times 10^{-3}}{800} = 5 \times 10^{-7}\text{m} = 5000\text{Å}$$

Ex. 45. A thin film of transparent material of refractive index 1.50 for the wavelength 5890 Å is inserted in one of the arms of Michelson Interferometer. A shift of 10 fringes is observed. Calculate the thickness of the film.

Sol. If a film of thickness t and refractive index μ is introduced in the path of one of the interfering beams then it produces an additional path difference of $2(\mu - 1)t$.If n is the number of fringes shifted, then,

$$2(\mu - 1)t = n\lambda$$

Or $$t = \frac{n\lambda}{2(\mu - 1)}$$

$$= \frac{10 \times 5890 \times 10^{-10}}{2(1.5 - 1)} = 5.89 \times 10^{-6}\text{m} = 5.89\mu\text{m}$$

Ex.46 Calculate the distance between successive positions of the movable mirror of Michelson interferometer giving best fringes in case of sodium source having wavelengths 5896 Å and 5890 Å.

Sol. Here $\lambda_1 = 5896$ Å $= 5.896 \times 10^{-7}$m and $\lambda_2 = 5890$ Å $= 5.89 \times 10^{-7}$m

The small difference in two wavelengths:

$$\Delta\lambda = \lambda_1 - \lambda_2 = \frac{\lambda_1\lambda_2}{2(x_2 - x_1)}$$

Or $$(x_2 - x_1) = \frac{\lambda_1\lambda_2}{2\,\Delta\lambda} = \frac{5.896 \times 10^{-7} \times 5.89 \times 10^{-7}}{2 \times (5.896 - 5.890) \times 10^{-7}}$$

Or $$(x_2 - x_1) = 28.94\text{Å}$$

*****Review Questions and Problems*****

Based On YDSE

1. Briefly outline the wave theory of light.
2. What is a wave front? How does it propagate?
3. What do you understand by phase difference and path difference?
4. Explain clearly Huygens' principle for the propagation of light.
5. Describe the Young's double slit experiment. What is its importance in physics?
6. Describe Young's double-slit experiment for obtaining interference fringes. Derive the expressions for (a) variation of intensity with phase difference, (b) positions of maxima and minima, and (c) fringe width.
7. Obtain an expression for fringe width in case of Young's double slit experiment. Prove that in this case of interference dark and bright bands are of equal width.
8. Young's double slits experiment is carried out with monochromatic light in air. What will be the change in wavelength and fringe width when the apparatus is immersed in water or the medium is replaced by an optically denser medium'?
9. Obtain the relations for constructive and destructive interference due to two slits clearly pointing out the conditions under which these equations are deduced.
10. Draw a schematic labelled diagram for the experimental determination of the wavelength of light by Young's interference experiment. Derive the formula used.
11. Justify that the formation of interference fringes in Young's experiment is in accordance with the conservation of energy.
12. Show that the distance between adjacent bright bands is inversely proportional to the distance between the slits.
13. Two coherent sources whose intensity ratio is 100: 1 produce interference fringes. Deduce the ratio of maximum intensity to minimum intensity in fringe system. [Ans. 121 : 81]
14. In an interference pattern with two coherent sources the amplitude of intensity variation is found to be 5% of the average intensity. Deduce the relative intensities of the interfering sources. [Ans. = 1600:1]
15. If in an interference pattern, the ratio between maximum and minimum intensities is 36:1, find the ratio between the amplitudes and intensities of the two interfering waves.
 [Ans. $a_1/a_2 = 7/5$, $I_1/I_2 = 49/25$]
16. At a certain point on a screen the path difference for the two interfering rays of same amplitude is $1/8^{th}$ of a wavelength. Find the ratio of intensity at this point to that at the centre of a bright fringe. [Ans. 0. 853]
17. A double slit arrangement produces interference fringes for sodium light (λ = 5800 Å) that are 0.20° apart. For what wavelength would the angular separation be 10% greater?
 [Ans. 6479 Å]

Based On Interference

18. What do you mean by interference of light?
19. What do you mean by coherent sources?
20. Explain why two independent sources can never be coherent.
21. Can two halves of a 100-Watt bulb be coherent?
22. What are the conditions of maxima and minima in an interference pattern?"

23.Name two broad methods for the production of coherent sources?
24.What are the conditions for interference of light?
25.What is meant by interference of light? State the superposition principle.
26.Explain the concept of coherence.
27. What are the necessary conditions for obtaining interference fringes?
28.What do you mean by constructive and destructive interference of light?
29.What are coherent sources? List down conditions for obtaining a good sustained interference pattern.
30.What are coherent sources? What are the conditions for two sources to be coherent? How are they realised in practice? Can two independent sources become coherent?
31.State and explain in brief the conditions for a) observance b) good contrast of fringes and c) stationary interference pattern.
32.Obtain an expression for the shift in the fringe pattern when a thin transparent slab is
33.Introduced in front of one of the slits in Young's experiment. How are coherent sources obtained in practice?
34.What are coherent sources? Why two independent sources of light cannot be coherent? What are the two general methods for obtaining coherent sources?

Based On Fresnel Biprism

36.What is a Fresnel's biprism?
37.Explain the formation of coherent sources by the use of a biprism.
38.Draw a labelled ray diagram depicting interference by a biprism.
39.What is biprism? Explain the construction and working of it with applications.
40.Discuss how coherent sources are produced with the help of Fresnel's biprism.
41.What happens when a very thin film of refracting material is placed in the path of one of the interfering waves?
42.What is the effect on interference pattern when monochromatic light in Fresnel's biprism experiment is replaced by white light?
43.Describe the geometrical features of Fresnel's biprism. How can it be used to find the wavelength of light?
44.Describe the construction, theory and working of Fresnel's biprism experiment to find the wavelength of light?
45.Define coherent and incoherent sources of light. Illustrate two methods for producing two coherent sources. Explain how the fringe-width of Fresnel biprism fringes can be increased?
46.Discuss conditions necessary for observing interference fringes of light. How are these satisfied in a biprism?
47.Discuss the conditions for sustained interference.
48.Describe Fresnel's biprism method for the determination of the wavelength of light.
49.Explain how are these fringes affected when (i) white light is used to illuminate the slits, (ii) the width of the slits is increased continuously and (iii) the edge of the biprism is not parallel to the slit.
50.(i) Interference pattern will consist of few coloured fringes on either side of central white fringe, (ii) the fringes become broader and the contrast between dark and bright fringes decreases. When the width of the slit become equal to half the fringe width, the fringes disappear, (iii) Lateral shift of fringes will occur]
51.Briefly discuss the effect of introducing a thin plate in the path of one of the interfering beams in a biprism. Deduce an expression for the displacement of the fringes. Show how the

direction of displacement is in favour of the wave theory of light. Show how this method is used for finding the thickness of a mica-sheet.

52. In a biprism experiment the micrometer readings for zero order and tenth order fringes are 1.25 mm. and 2.37 mm, respectively when light of $\lambda = 6.0 \times 10^{-5}$ cm. is used. What will be the positions of zero order and tenth order fringes if λ is changed to 7.5×10^{-5} cm.? [Ans. 1.25 mm., 2.65 mm.]
53. The inclined faces of glass biprism ($\mu = 1.5$) make an angle of 2° with the base of the prism. The slit is 10 cm. away from the prism and is illuminated by light of wavelength 5500 Å. Calculate (i) the separation between the coherent sources formed by the biprism, and (ii) the fringe width at a distance of 1 metre from the slit. [Ans. (i) 0.349 cm. (ii) 0.0157 cm.]
54. Fresnel biprism of angle 1° and refractive index 1.56 forms interference fringes on a screen placed 80 cm. from the biprism. If the distance between the slit and biprism is 30 cm. find the fringe separation when the wavelength of light used is 5000Å. Also calculate the fringe separation when distances 80 cm. and 30 cm. above are interchanged. [Ans. ω = 0. 009375 cm., 0. 0031 cm.]
55. A biprism of obtuse angles 176° is made of glass of refractive index 1.5. A slit illuminated with monochromatic light is placed 20 cm behind and the width of interference fringes formed on a screen 80 cm. in front of biprism is found to be 8.25×10^{-3} cm. Calculate the wavelength of light. [Ans. 5.757×10^{-5} cm.]
56. In a biprism experiment the eye-piece was placed at distance of 120 cm. from the source. The distance between two virtual images was found equal to 0.075 cm. Find the wavelength of light source, if eye-piece micrometer is moved through a distance 1.888 cm. for 20 fringes cross the field of view. [Ans. 5900 Å]
57. The distance between the slit and biprism and that between the biprism and screen are each 50 cm. The obtuse angle of biprism is 179° and its refractive index is 1.5. If the width of the fringes is 0.0135 cm. calculate the wavelength of light.
58. In a two-slit experiment with monochromatic light fringes are observed on a screen placed at some distance from the slits. If the screen is moved by 5×10^{-2} m towards the slits, the change in fringe width is 3×10^{-3} m. If the distance between the slits is 10^{-3} m. Calculate the wavelength of light used. [Ans. 6000 Å]
59. The distance between the two images for the two positions of the lens were respectively 0.3 and 1.2 mm and the width of 10 fringes was 9.720 mm. Calculate (i) the distance between the focal plane of the eye-piece and the plane of interfering [Ans. (i) 99 cm., (ii) 5891 Å]
60. In a two-slit experiment with monochromatic light fringes are observed on a screen placed at some distance from the slits. If the screen is moved by 5×10^{-2} m towards the slits, the change in fringe width is 3×10^{-3} m. If the distance between the slits is 10^{-3} m. calculate the wavelength of light used. [Ans. 6000 Å]
61. If the angle of biprism is 0.01 radian and the slit is placed at a distance of 10 cm. from the biprism and the eyepiece is kept at a distance of 90 cm from the biprism. Find the fringe width when the wavelength of light used is 6000 Å and the refractive index of glass is 1. 5. [Ans. 6×10^{-4} m.]
62. A Fresnel's biprism with apex angle 1°30' is used to form interference fringes. Its refractive index is 1.52. Calculate the fringe width for wavelength of light $\lambda = 6563$ Å, when the distance between the source and biprism is 20 cm. and that between the biprism and screen is 80 cm. [Ans. 0.0121 cm.]

63.In a biprism arrangement 40 fringes are seen in the field of view when light of wavelength 5893 Å is used. If light of wavelength 4358 Å be used, how many fringes will be seen in the same field of view? [Ans 54]

64.In an interference pattern, at a point we observe the 12th order maximum for wavelength $\lambda_1 = 6000$ Å. What order will be visible here if the source is replaced by light of wavelength $\lambda_2 = 4800$ Å. [Ans 15]

65.Interference fringes are produced by Fresnel's biprism in the focal plane of an eye-piece 2 m. away from the slit the two images of the slit that are formed for each of the two positions a convex lens placed between the biprism and the eyepiece are found to the separated by 4.5 mm. and 2.9 mm. respectively. If the width of the interfering fringes be 0.326 mm, find the wavelength of the light used.

66.In biprism experiment using mercury green light (λ = 5461 Å), a thin transparent film placed in the path of one of the interfering beams causes a shift of 5 fringes. If the refractive index of the film material is 1.33, calculate the thickness of the film. [Ans. 0.827 mm.]

67.In a biprism experiment the distance between the slit and the screen in 180 cm. The biprism is 60 cm away from the slit and its refractive index is 1. 52. When a source of wavelength 5893 Å is used, the fringe width is found to be 0.010 cm. Find the angle between the two refracting surfaces of the biprism. [Ans. $\approx 1^\circ$]

68.A double slit of separation 1.5 mm. is illuminated by white light (between 4000 Å to 8000 Å). On a screen 120 cm. away coloured interference pattern is formed. If a pin hole is made on this screen at a distance of 3 mm. from the central white fringe, what wavelengths will be absent in the transmitted light? [Ans. 6818.2 Å, 5769.2 Å, 5000 Å, 4411.9 Å]

69.Fresnel's biprism fringes are observed with white light when a thin transparent sheet covers one-half of the biprism, the central fringe shifts sideways by 14.97 mm. With the same geometry, the fringe width with mercury green light (5461Å) comes 0.274 mm. Deduce the thickness of the sheet, assuming the refractive index of its material 1.58.[Ans. 5.14 mm.]

70.Fresnel's fringes are produced with homogeneous light of wavelength 6×10^{-5} cm. A thin glass film = 1.5) is interposed in the path of one of the interfering beams. The central bright band is shifted to the position previously occupied by the fifth bright fringe. Find the thickness of the film. [Ans. 6×10^{-4} cm.]

71.If fringe width with wavelength of light $\lambda = 5.89 \times 10^{-7}$ metre is 0.431 mm. and shift of white central fringe on introducing a mica sheet in one path is 1.89 mm., calculate the thickness of the mica sheet (μ= 1.591). [Ans. 4.38×10^{-7} metre]

72.On introducing a thin sheet of mica (thickness 12×10^{-5} cm) in the path of one of the interfering beams in a biprism experiment, the central fringe is shifted through a distance equal to the spacing between successive bright fringes. Calculate the refractive index of mica. ($\lambda \times 10^{-7}$ m) [Ans. 1.5]

73.A thin sheet of glass (μ = 1.5) of 6 microns thickness introduced in the path of one of the interfering beams in a biprism arrangement shifts the central fringe to a position normally occupied by the fifth fringe. Find the wavelength of light used. (1 micron = 10^{-6} m) [Ans. 6000 Å]

74.A thin mica sheet (μ = 1.6) of 7-micron thickness introduced in the path of one of the interfering beams in biprism arrangement shifts the central fringe to a position previously occupied by the seventh bright fringe from the centre. Calculate the wavelength of light used. [Ans. 6000 Å] .

75.A glass plate 12×10^{-3} mm. thick is placed in the path of one of the interfering beams in a biprism arrangement using monochromatic light of wavelength 6000 Å. If the central band shifts a distance equal to width of 10 bands, find the refractive index of glass. What is the thickness of the plate of diamond of refractive index 2.5 that has to be introduced in the path of second beam to bring the central band to its original position? [Ans. $\mu = 1.5$, $t = 4 \times 10^{-4}$ cm]

76. On introducing a thin sheet of glass ($\mu = 1.5$) in one of the two interfering beams in a biprism experiment with white light, the shift in central fringe is observed 4.5 mm. With sodium light ($\lambda = 5893$ Å) the same arrangement gives fringe width equal to 0.225 mm. Calculate the thickness of the glass. [Ans. 2.357×10^{-3} cm.] [Hint. Shift $y_o = 4.5$ mm, fringe width (for $\lambda = 5893$ Å) is 0.225 mm. Now the relation $(\mu - 1)\, t = n\lambda$ may be used to find t]

77. In double slit arrangement fringes are produced using light of wavelength 4800 Å. One slit is covered by a thin plate of glass of refractive index 1.4 and the other slit by another plate of glass of same thickness but of refractive index 1.7. On doing so the central bright fringe shifts to the position originally occupied by the fifth bright fringe from the centre. Find the thickness of the glass plate. [Ans. 8.0×10^{-6} m]

Based on Thin films

78. Explain formation of colours when white light is incident on a transparent thin film.
79. Explain why "A thick film shows no colours in reflected white light".
80. Explain why "An excessively thin film seen in reflected white light appears perfectly black".
81. Explain the need of extended source in interference with division of amplitude.
82. Discuss the phenomenon of interference of light in thin films and obtain the conditions of maxima and minima for the reflected light.
83. Give the nature of fringes obtained in thin parallel films.
84. Discuss the interference from parallel thin film. Describe salient features of the fringes formed. Give one of its applications.
85. Discuss the phenomenon of interference of light in thin films and obtain the conditions of maxima and minima. Show that the interference patterns in reflected and transmitted lights are complimentary.
86. What will happen if a wedge-shaped film is placed in white light?
87. Using an optical method how would you determine the thickness of a piece of transparent cello tape? Explain.
88. Describe the interference observed when a thin parallel shaped film is seen by reflected light normally.
89. Explain how interference fringes are formed by a thin wedge-shaped film when examined by normally reflected light. Find the expression for fringe-width. How will you estimate the difference of film thickness between two points?
90. Light of wavelength 4500 Å falls on a sheet of transparent material whose index of refraction is 1.50. What must be the minimum thickness of the sheet in order that there shall be constructive interference in the reflected beam? [Ans. 750 Å]
91. A parallel beam of light of wavelength 5890 Å is incident on a thin glass plate of refractive index 1.5 such that angle of refraction into plate is 60°. Calculate the smallest thickness of the plate which will make it appear dark by reflection. [Ans. 3927 Å]
92. A beam of parallel rays is incident at an angle of 30° with the normal on a plane parallel film of thickness 0.00004 cm. and refractive index 1.5. Show that the reflected light whose wavelength is 7.539×10^{-5}cm will be strengthened by interference.
93. White light is incident on two parallel glass plates separated by an air film of thickness 0.001 cm and the reflected light is examined by a spectroscope. Find the number of dark bands seen in the spectrum between the wavelengths 4×10^{-5} cm and 7×10^{-5} cm when light incident at an angle of 30° to the normal to the surface. [Ans. 19]

94.White light reflected at perpendicular incidence from a soap film, has in the visible spectrum, an interference maximum at 6000 Å and a minimum at 4500 Å with no minimum in between. If μ = 1.33 for the film, what is the thickness of soap film, assumed uniform. [Ans. 3383 Å]

95.A soap film is seen in sodium yellow light (λ = 5893 Å) with rays falling at 30° from normal. Bright fringes occur, at two points A and B with four more bright fringes in between. Deduce the difference of thickness of the film between these points, μ of soap-film = 1.33. [Ans. 1.19×10^{-6} m.]

96.A soap film of refractive index 4/3 and of thickness 1.5×10^{-6} m., is illuminated by white light incident at an angle of 45°. The light reflected by it is examined by a spectroscope in which it is found a dark band corresponding to a wavelength of 5×10^{-7} m. Calculate the order of the interference band. [Ans. 7]

97.Find the thickness of a wedge-shaped air film at a point where fourth bright fringe is situated. Wavelength of sodium light = 5893 Å [Ans. 1.02×10^{-6} m.]

98.Two glass plates enclose a wedge-shaped air film, touching at one edge and are separated by a wire of 0.03 mm. diameter at a distance of 15 cm. from the edge. Monochromatic light, λ = 6000 Å from a broad source fall normally on the film. Calculate the fringe width of the fringes thus formed. [Ans. 0.15 cm.]

99.Interference fringes are produced by monochromatic light falling normally on a wedge-shaped film of cellophane whose refractive index is 1.4. The angle of wedge is 20 sec. and the distance between successive fringes is 0.25 cm. Calculate the wavelength of light. [Ans. 7×10^{-5} cm.]

100.Two optically flat discs lie one on the other. A sheet of paper 0.500 mm. thick is inserted between the two discs at one edge. How many dark interference fringes will appear by reflected light if the discs are illuminated by light of wavelength 5890 Å? What will be the shape of the fringes? [Ans. 171]

101.A wedge-shaped air film, having an angle of 40 seconds, is illuminated by monochromatic light and fringes are observed vertically through a microscope. The distance measured between the consecutive fringes is 0.12 cm. Calculate the wavelength of light used. [Ans. 4656 Å]

102.Light of wavelength 5500 Å falls normally on a thin, wedge-shaped film of refractive index 1.4 forming fringes that are 2.5 mm. apart. Find the angle of wedge in seconds. [Ans. 16.2 seconds]

103.If the angle of wedge is 0.25 degree of arc and the wavelength of sodium D lines are 5890 Å are 5896 Å, find the distance from the apex of the wedge at which the maxima due to each wavelength first coincide when observed in reflected light. [Ans. 6.63 cm.]

104.Two plane rectangular pieces of glass are in contact at one edge and separated by a hair at opposite edge so that a wedge is formed. When light of wavelength 6000 Å falls normally on the wedge, nine interference fringes are observed. What is the thickness of the hair? [Ans. 2.7×10^{-6} m.]

105.Light of wavelength 5.9×10^{-7} m is incident on a soap film at 30°; dark bands are observed 5.00 mm apart. If μ = 1.33, find the angle between the faces of the film. [Ans: 10°]

Based on Newton Rings

106.What are Newton's rings?

107.Explain why Newton's rings are circular.

108.Explain why a narrow source is used in the biprism experiment, whereas an extended source is used in Newton's rings experiment.

109.What are Newton's rings? Explain the formation of Newton's rings by reflected system of light. Also show that spacing between rings goes on decreasing with increased order.

110.Write a note on Newton's ring. What determines whether the centre shall be bright or dark?

111.With the help of a labelled ray diagram discuss the formation of Newton's rings by reflected light. Hence derive an expression for the diameter of n th dark ring. Explain the formation of fringes in Newton's ring experiment. Give its application to find out wavelength of light.

112.With the help of a neat diagram show an experimental arrangement to produce Newton's rings in reflected sodium light. Prove that in reflected light the diameter of the dark rings is proportional to the square root of the natural number.

113.Prove that the diameters of the bright rings are proportional to the square root of odd natural numbers.

114.Explain with the help of a diagram how Newton's rings are formed. Describe how Newton's rings can be used to determine the wavelength of light.

115.Describe and explain formation of Newton's ring in reflected monochromatic light. Prove that in reflected light diameters of the dark rings are proportional to the square root of natural numbers.

116.Discuss in brief the conditions under which the centre of Newton's ring is bright or dark.

117.What are the conditions for maxima and minima in case of Newton's ring due to reflected light? How the refractive index of any liquid can be determined by Newton's ring

118.Explain fringe width obtained in Newton's rings experiment. Derive an expression for the radius of curvature in the Newton's ring experiment?

119.Explain the phenomenon of interference in thin film and also explain with theory

120.Describe Newton's rings method for determining the refractive index of a liquid. Derive the formula used.

121.Why does the centre of Newton's ring appear dark in reflected light?

122.Explain why Newton's rings are circular but air- wedge fringes are straight.

123.What change would you expect in interference pattern of Newton's rings if transparent plate below the lens is replaced by a plane mirror?

124.What change will occur in interference pattern when a little water is introduced between the lens and plate in Newton's rings arrangement?

125.What are Newton's rings? Explain how they can be used to find wavelength of light.

126.Explain why a narrow source is necessary for biprism experiment but an extended source is required for Newton's rings experiment.

127.Explain with theory Newton's rings experiment to determine the wavelength of monochromatic light and discuss which source is preferred: point or extended.

128.A planoconvex lens of radius 3 m is placed on an optically glass plate and is illuminated by monochromatic light. The diameter of 8th dark ring in the transmitted system is 0.72 cm. Calculate the wavelength of light used. Ans. 5760 Å]

129.In Newton's rings arrangement the diameter of nth and (n + 14) th rings are 4.2 mm. and 7.0 mm respectively. Radius of curvature of planoconvex lens is 1 m. Calculate the wavelength of light. [Ans. 5. 6 × 10^{-7} m]

130.A convex lens of radius 3.50 m placed on a flat plate and illuminated by monochromatic light gives the 6th bright ring of diameter 0.68 cm. Calculate the wavelength of light used. [Ans. 6000 Å]

131.Newton's rings are made with light of wavelength 6400 Å and a thin layer of oil (μ= 1. 60) formed between the curved surface of a piano-convex lens (radius of curvature = 80 cm, μ= 1. 65) and a plane glass plate (μ = 1.55). Calculate the radius of the smallest dark ring.Ans. 0.04 cm.]

132.In an arrangement for observing Newton's rings with two different media between the glass surfaces, the n^{th} rings have diameters as 10: 7. Find the ratio of the refractive indices of the two media. [Ans. 49: 100]

133.In the experiment of Newton's rings, the diameter of the fifth dark ring is measured as 3.06 mm. Calculate the radius of the 15^{th} dark ring.

134.If a small amount of water is poured between the lens and glass plate, calculate the radii of 5^{th} and 15^{th} dark rings. Given refractive index of water = 1.33.[Ans. (i) 5.3 mm., (ii) 2.653 mm, 4.596 mm.]

135.A convex lens is placed on a slab of plane glass and is illuminated by monochromatic light. The diameter of the 10^{th} dark ring is measured in reflected light and is found to be 0.433 cm. Find the wavelength of light, if the radius of curvature of the lower face of the lens is 70 cm. [Ans. 6695 Å]

136.Newton's rings are observed normally in the reflected light of wavelength 5893 Å. The diameter of the 10^{th} dark ring is 0.005 m. Find the radius of curvature of the lens and the thickness of the air film. [Ans. R = 1.062 m., t = 2.946×10^{-6} m.]

137.In a Newton's rings experiment, the diameter of 5^{th} dark ring is reduced to half of its value after introducing a liquid below the convex surface. Calculate the refractive index of liquid. [Ans. 4]

Based on Michelson Interferometer

138.The Michelson Interferometer is based upon the interference of light due to division of wavefront or division of amplitude. Justify your answer.

139. Explain why a compensating plate is needed in Michelson Interferometer?

140. Explain the formation of fringes in Michelson Interferometer with a suitable diagram.

141.Explain the formation of fringes in Michelson's interferometer. Give its application to determine the wavelength of light.

142. When a thin film of transparent material of refractive index 1.45 for $\lambda = 5890 \times 10^{-8}$ cm is inserted in one of the arms of a Michelson interferometer, a shift of 65 circular fringes is observed. Calculate the thickness of the film. [Ans. t = 0.00425 cm].

Diffraction: Fraunhofer and Fresnel diffraction, Fraunhofer diffraction for single slit, double slit and N slit (diffraction grating), Fraunhofer diffraction from circular aperture, Rayleigh criterion, Resolving power of optical instruments.

Chapter 4

Diffraction of Light

Introduction

If an object with a sharp edge is placed in between a light source and the screen, the image of an object is sharp except at the edges. At the edges, the intensity of light is redistributed and the bright side contains some darkness and the dark side contains some brightness. It seems as if light bends at the sharp edges. This effect is more pronounced if the size of the object is of the order of the wavelength of light. This phenomenon of light is termed as diffraction of light. In other words, "the departure from the rectilinear path of light, when it meets the objects of the sizes comparable of the wavelength of light is called diffraction of light".

Satisfactory explanation of diffraction was given by A. J. Fresnel in 1815. He combined Huygen's secondary wavelets theory with the interference of light. He suggested that the diffraction is the interference of light coming from a number of Huygen's secondary wavelets from the same source. According to the theory of secondary sources, when light interacts at an interface with a medium, number of Huygen's secondary sources are generated at the point of interaction with same frequency as the original light. If the point of interaction is an object whose size is of the order of wavelength of light, the secondary sources work as spherical sources and emits light in different direction whose intensity is according to Fresnel's equation i.e. proportional to $(1 + \cos\theta)$, where θ is an angle from the perpendicular direction to the surface. Now, if the size of a slit or an obstacle is of the order of wavelength of light, the entire slit or say an obstacle will produce a number of Huygen's secondary sources and the superposition of light wave trains from these sources, reaching at a point on the screen will produce the so-called diffraction phenomena. So, if a wavefront of light is exposed to a slit or an obstacle of the size of the order of wavelength of light, the superposition of secondary wavelets originating from the exposed part of the same wave front will produce diffraction effect on the screen.

Here, the only conceptual difference between interference and diffraction can be understood as follows;

Interference: It is the modification of intensity of light at the exposed part of the screen, due to superposition of light from two difference sources. Obviously, to see the sustained and well-defined pattern, the two sources must be coherent and of nearly equal intensity.

Diffraction: It is the modification of intensity of light at the screen, due to superposition of a number of secondary wavelets originated from the exposed part of the same wave front at the slit or an obstacle.

Fresnel and Fraunhofer Diffraction

As light produced from a source or any object is exposed to light, the Huygen's secondary sources starts interfering each other. Up-to a certain distance, the variation of intensity varies in a haphazard way. The explanation to such kind of pattern is given by Fresnel, called as Fresnel diffraction. After a certain distance, the variation of intensity due to superposition of waves coming out from a single source varies in different manner.

When light originates from any source, up to some distance called Fresnel distance, the intensity of light varies in one way and after that the light varies in another way. The analysis of diffraction pattern in the near field of the source is done by Fresnel and in the far field the treatment is done by Fraunhofer. These two treatments of the diffraction portions of the complete diffraction pattern are termed as Fresnel and Fraunhofer diffraction respectively. In a clear and separate way, the diffraction can also be defined as;

4.1 Fresnel Diffraction

If the slit or obstacle is exposed with spherical wave front, the phenomena is explained by Fresnel's method, called Fresnel's type of diffraction. In other words, "If the source, slit and screen are at finite distance, the pattern seen on the screen is termed as Fresnel diffraction".

4.2 Fraunhofer Diffraction

If the slit is exposed to a plane wave front, the phenomenon is called Fraunhofer type of diffraction. Actually, if the slit and screen are near to a point source, the wave front will be considered as spherical and the diffraction is called Fresnel's type. And if slit and screen are placed at infinite (very far) distance from the point source, the wave front may be considered as plane and the pattern obtained can be understood by Fraunhofer's explanation, called Fraunhofer type of diffraction.

Here, we will discuss only Fraunhofer diffraction as prescribed in the syllabus of the university. We will explain this phenomenon in case of single slit, two slits and multiple slits. Although the two slit diffraction phenomenon is not in the syllabus, but the discussion will give clear vision about when the two slits are exposed to light there will be a two slit Fraunhofer diffraction and or Young's two slit interference experiment.

4.3 Fraunhofer Diffraction at Single Slit

Let a slit of width d, of the order of wavelength of light, is exposed to a plane wave front of light and the pattern is obtained on the screen placed at a far distance as shown in Fig.4.1. This whole arrangement of set up is placed in such a way that it fulfills the required conditions for obtaining Fraunhofer type of diffraction. Now according to the definition of diffraction of light, the phenomena obtained on the screen can be explained by considering the superposition of a number of secondary wavelets coming from the Huygen's secondary sources which are generated due to the exposed part of the plane wave front through the slit.

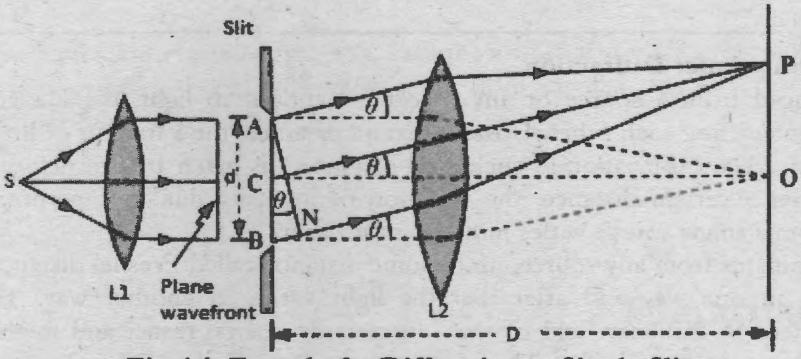


Fig.4.1. Fraunhofer Diffraction at Single Slit

Now according to the Huygen's theory of secondary sources, let there are n number of secondary sources within the slit each with width d. So, the distance between the two successive sources may be taken as $\frac{d}{n}$. Now it may be considered that there are n numbers of secondary waves reaching at a point P on the screen, each having nearly equal amplitude and path difference between two successive waves as:

Path difference = $\frac{d}{n}\sin\theta$,

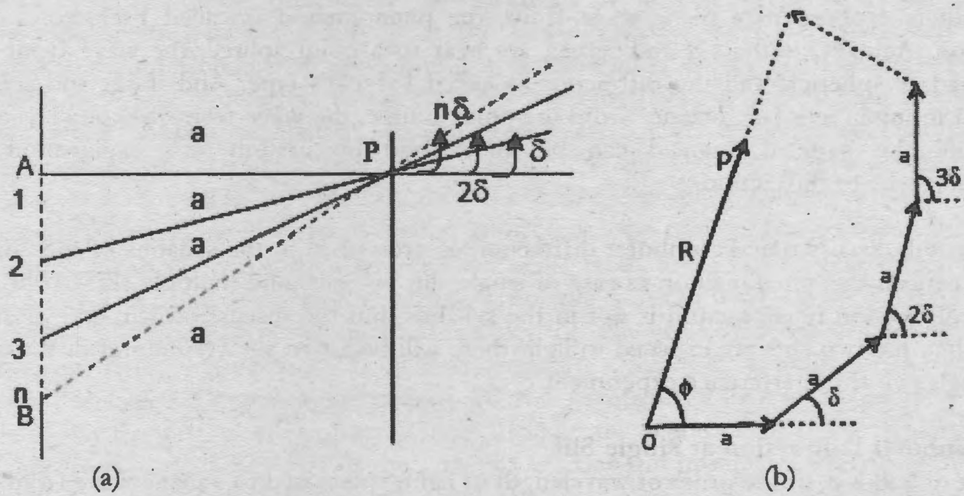


Fig.4.2. Composition of n simple harmonic motions of equal amplitudes, periods and phases increasing in arithmetic progression

Where θ is the direction of point P from the horizontal line from the centre of the slit. As the slit width is very small i.e., comparable to the wavelength of light, the direction of each wave from the Huygen's sources, reaching at P with the horizontal may be taken as θ.

So, phase difference between two successive waves will be:

Phase difference, $\delta = \frac{2\pi}{\lambda}\left(\frac{d}{n}\sin\theta\right)$

Now, we have to find out the resultant amplitude/intensity at any arbitrary point P on the screen due to n number of wave vectors each having amplitude a and constant phase difference between two consecutive wave vectors as δ.

According to the Fig. 4.2, let R be the resultant of these n numbers of wave vectors at P having phase ϕ.Then, resolving R into horizontal and vertical components, we have:

$$R\cos\phi = a + a\cos\delta + a\cos 2\delta + a\cos 3\delta + \ldots\ldots\ldots a\cos(n-1)\delta \qquad (1)$$

$$R\sin\phi = 0 + a\sin\delta + a\sin 2\delta + a\sin 3\delta + \ldots\ldots\ldots\ldots a\sin(n-1)\delta \qquad (2)$$

Multiplying equation (1) by $2\sin\frac{\delta}{2}$.

$$R\,2\sin\frac{\delta}{2}\cos\phi = 2a\sin\frac{\delta}{2} + 2a\sin\frac{\delta}{2}\cos\delta + 2a\sin\frac{\delta}{2}\cos 2\delta + a\sin\frac{\delta}{2}\cos 3\delta + \text{-----------------} a\sin\frac{\delta}{2}\cos(n-1)\delta \qquad (3)$$

Using trigonometric identity: $2\sin A\cos B = \sin(A+B) - \cos(A-B)$,the equation (3) takes the form;

$$R\ 2\sin\frac{\delta}{2}\cos\phi = 2a\sin\frac{\delta}{2} + a\sin\frac{3\delta}{2} - a\sin\frac{\delta}{2} + a\sin\frac{5\delta}{2} - a\sin\frac{3\delta}{2}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots a\sin\frac{(2n-1)\delta}{2} - a\sin\frac{(2n-3)\delta}{2}$$

$$= a\left[\sin\frac{\delta}{2} + \sin\frac{(2n-1)\delta}{2}\right]$$

$$= 2a\sin\frac{n\delta}{2}\cos\frac{(n-1)\delta}{2}$$

$$\text{Or } R\cos\phi = \frac{a\sin\frac{n\delta}{2}\cos\frac{(n-1)\delta}{2}}{\sin\frac{\delta}{2}} \qquad (4)$$

Similarly, multiplying equation (2) by $2\sin\frac{\delta}{2}$ and after simplification, we get

$$R\sin\phi = \frac{a\sin\frac{n\delta}{2}\sin\frac{(n-1)\delta}{2}}{\sin\frac{\delta}{2}} \qquad (5)$$

Squarring and adding, equation (4) and (5) and the taking the square root we get:

$$R = \frac{a\sin\frac{n\delta}{2}}{\sin\frac{\delta}{2}} \qquad (6)$$

The above result can always be used as a standard result to find the resultant amplitude of n number of waves each having amplitude 'a' and phase difference between consecutive waves as δ.

Dividing (5) by (4), we get;

$$\tan\phi = \tan\frac{(n-1)\delta}{2}$$

Or, the final phase, $\phi = \frac{(n-1)\delta}{2}$

In nutshell, if we want to add 'n' number of waves each having amplitude 'a' and constant phase difference δ between the two consecutive waves, then after superposition, the resultant of all the waves, can be written as

$$R = \frac{\text{amplitude of a wave} \times \sin \frac{(\text{number of waves} \times \text{constant phase difference between two consecutive waves})}{2}}{\sin \left(\frac{\text{constant phase difference between two consecutive waves}}{2} \right)} \quad (7)$$

Now, putting the values of amplitude and phase difference for the particular case, we have;

$$R = \frac{a \sin \frac{n}{2} \left(\frac{2\pi d}{\lambda n} \sin \theta \right)}{\sin \frac{\left(\frac{2\pi d}{\lambda n} \sin \theta \right)}{2}}$$

$$R = \frac{a \sin \frac{\pi d}{\lambda} \sin \theta}{\sin \frac{\pi d}{\lambda n} \sin \theta}$$

Let, $\alpha = \frac{\pi d}{\lambda} \sin \theta$, then

$$R = \frac{a \sin \alpha}{\sin \frac{\alpha}{n}},$$

If there are a large number of Huygen's sources, i.e., for large value of n, we have $\frac{\alpha}{n}$ very small, then

$$R = \frac{a \sin \alpha}{\frac{\alpha}{n}}$$

Or $\quad R = \frac{n a \sin \alpha}{\alpha} = \frac{A \sin \alpha}{\alpha}$,

Where A = na is the sum of the amplitudes of all the waves or we can say amplitude of all the waves when they are in the same phase.

In other words, we can say:

Resultant R = Sum of all the waves in same phase $\times \frac{\sin \alpha}{\alpha}$,

Where $\quad \alpha = \frac{\pi}{\lambda} \times$ path difference between two extreme waves in the slit.

And $\quad$ Phase $\phi = \frac{(n-1)\delta}{2} = \frac{n\delta}{2}$

i.e., Phase of resultant wave is $= \frac{n}{2} \times$ Phase between two successive waves.

Now, let us discuss the conditions for maximum and minimum intensity at the point P. As the intensity is proportional to the square of the amplitude, taking proportionality constant one, the intensity at any point in the diffraction pattern obtained on the screen is given by;

$$I = R^2 = A^2 \frac{\sin^2 \alpha}{\alpha^2} \quad (8)$$

The equation (8) gives the variation in the intensity in the diffraction pattern, depending upon $\alpha = \frac{\pi}{\lambda} d \sin \theta$.

Conditions for maximum and minimum intensity can be find out equating $\frac{d(I)}{d\alpha} = 0$, so differentiating equation (8) we have:

$$\frac{d(I)}{d\alpha} = \frac{d}{d\alpha} \left(A^2 \frac{\sin^2 \alpha}{\alpha^2} \right) = A^2 \frac{2 \sin \alpha}{\alpha} \left(\frac{\alpha \cos \alpha - \sin \alpha}{\alpha^2} \right) = 0,$$

$$\Rightarrow \frac{\sin \alpha}{\alpha} = 0 \quad \text{or} \quad \frac{\alpha \cos \alpha - \sin \alpha}{\alpha^2} = 0$$

$$\Rightarrow \frac{\sin \alpha}{\alpha} = 0 \quad \text{or} \quad (\alpha \cos \alpha - \sin \alpha) = 0$$

$$\Rightarrow \frac{\sin\alpha}{\alpha} = 0 \qquad \text{(a)}$$

$$\text{or} \quad \alpha = \tan\alpha \qquad \text{(b)} \qquad (9)$$

These two conditions 9(a) and 9(b) will decide the directions of maximum or minimum intensities in the diffraction pattern obtained in case of a single slit.

4.3.1 Condition for Minimum Intensity

If $\frac{\sin\alpha}{\alpha} = 0$, it is clear from equation (8) that the intensity will be zero.
So, the condition for minimum intensity,

$$\frac{\sin\alpha}{\alpha} = 0$$

$$\Rightarrow \sin\alpha = 0$$

But $\alpha \neq 0$, because if $\alpha = 0$, then $\frac{\sin 0}{0}$ will be indeterminate form and becomes as 1.

Thus, $\sin\alpha = 0$

$$\Rightarrow \alpha = n\pi, \quad \text{where } n = 1,2,3\ldots\ldots\ldots\ldots \text{ etc.}$$

$$\Rightarrow \alpha = \frac{\pi d}{\lambda}\sin\theta = \pm\, n\pi$$

$$\text{Or} \quad d\sin\theta = \pm\, n\lambda \qquad (10)$$

$$\Rightarrow \theta = \sin^{-1}\left(\pm\frac{n\lambda}{d}\right)$$

Equation (10) gives the directions of minimum intensities in the diffraction pattern.

4.3.2 Condition for Maximum Intensity

If $\alpha = \tan\alpha$, and plotting this equation as $y = \alpha$, and $y = \tan\alpha$, the points of intersection of $\tan\alpha$ curve and the line $y = \alpha$, will give all the points where the intensity is maximum as shown in Fig.4.3. As can be seen from the graph, first value of α is exact zero but the other values are nearly $\frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$ etc.

Exactly $\alpha = 0$, 1.430π, $2.46\,\pi$, 3.47π, …………………………………………… etc.

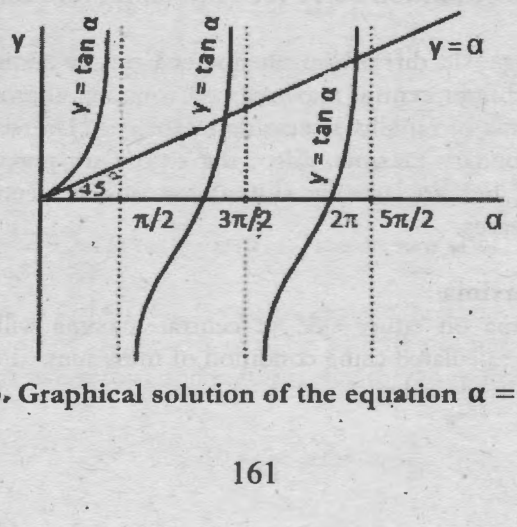


Fig 4.3. Graphical solution of the equation $\alpha = \tan\alpha$

4.3.3 Intensities of Successive Maxima in the Diffraction Pattern

(i) Zero order principle maximum

$\alpha = 0$, i.e., $\frac{\sin 0}{0} = 1$, by L' Hospital's rule, and the intensity is

$I_0 = A^2$

(ii) First, second, third, etc. orders principle maxima on either side

Taking $\alpha = \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$ etc., we have

$$I_1 = R^2 = A^2 \frac{\sin^2(\frac{3\pi}{2})}{(\frac{3\pi}{2})^2} = A^2 . \frac{4}{9\pi^2}$$

$$I_2 = R^2 = A^2 \frac{\sin^2(\frac{5\pi}{2})}{(\frac{5\pi}{2})^2} = A^2 . \frac{4}{25\pi^2}$$

$$I_3 = R^2 = A^2 \frac{\sin^2(\frac{7\pi}{2})}{(\frac{7\pi}{2})^2} = A^2 . \frac{4}{49\pi^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{etc.}$$

The ratio of successive maxima are given as:

$$1 : \frac{4}{9\pi^2} : \frac{4}{25\pi^2} : \frac{4}{49\pi^2} \quad \ldots\ldots \text{ etc.} \tag{11}$$

To put in a nutshell, the intensity decreases very rapidly as we move from central maximum to other principle maxima on either side or we can say maximum intensity is contained in the zero order principle maximum. It can also be seen from the conditions of maxima and minima that all the minima are equally spaced but maxima are not. The intensity curve is seen in Fig.4.4.

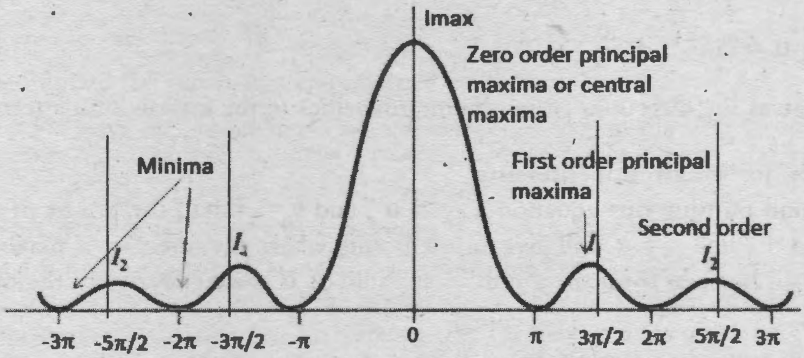


Fig.4.4. Intensity distribution curve for Fraunhofer diffraction at a single slit

Thus, the treatment of single slit diffraction phenomena can be summarized as: the diffraction pattern consists of a very bright central maximum surrounded alternatively by minima of zero intensities and feeble maxima of rapidly decreasing intensities. The width of the central maxima is twice the width of secondary maxima. Also, the secondary maxima do not fall exactly in between the two minima, but are actually shifted towards the centre of the pattern which decreases as the order increases.

4.3.4 Width of Central Maxima

Width between first minima on either side of central maxima will be the width of central maximum. So, it should be calculated using condition of minimum.

Equation of first minimum is given by;

$d \sin\theta = \pm\lambda$

Or $\theta = \pm \sin^{-1}(\frac{\lambda}{d})$, thus the central maximum extends from $+ \sin^{-1}(\frac{\lambda}{d})$ to $- \sin^{-1}(\frac{\lambda}{d})$

In other words, 2θ is called the angular width of the central maximum i.e.

Angular width of central maximum is $2\theta = 2[\pm \sin^{-1}\left(\frac{\lambda}{d}\right)]$

For a special case, if slit width d is equal to the wavelength of light used, what will happen to the diffraction phenomena? For this, let us start with the equation of minimum in the diffraction through a single slit.

As, $d \sin\theta = \pm\ n\lambda$

$$\theta = \pm \sin^{-1}(\frac{n\lambda}{d})$$

$$= \pm \sin^{-1}(n) \qquad (\text{as } \lambda = d)$$

The minimum value of n i.e.1 can be possible here for which the equation is satisfied. So

$$\theta = \pm \sin^{-1}(1)$$

$$\theta = \pm\frac{\pi}{2}$$

This implies that the central maximum will spread within 180^0 on the screen and no diffraction phenomena in this case will be observed. But this is not true. Actually, the complete brightness on the screen is the central maximum of diffraction phenomena, depicting that there is no diffraction phenomena occurring in this case. The decrement of intensity from centre of maxima with increase of angle with the central maximum, is due to the diffraction effect.

4.4 Fraunhofer Diffraction at Two Slits

Let there are two slit AB and CD Fig. 4.5. Let both are of exactly equal width i.e., AB = CD = d and the spacing between the BC is equal to e. This set of two slit is illuminated with a source of light of parallel wavefront with screen at infinity, as it is the pre-condition for Fraunhofer diffraction to occur. Now let us examine the intensity variation obtained at the screen.

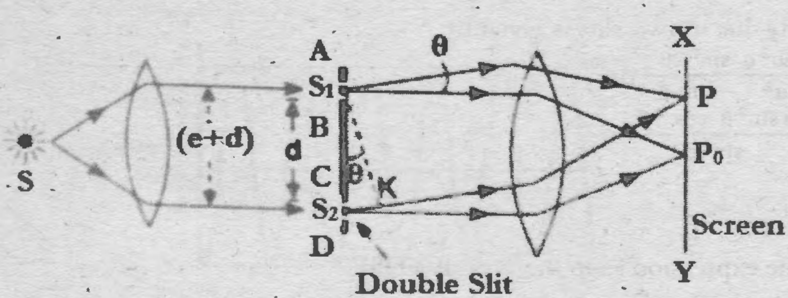


Fig. 4.5 Fraunhoffer Diffraction at two slits

Let P is a point of observation at the screen. It can be well understood that the intensity at P is the resultant of the superposition of amplitudes of lights at P due to individual slits AB and CD. This is nothing but resultant of two diffraction patterns due to two individual single slits AB and

CD. So far, we have already seen the treatment of Fraunhofer diffraction at single slit. According to that, we can take the resultant amplitude at P due to each slit, AB and CD which can be written as;

$$R = \frac{\text{amplitude of a wave} \times \sin \frac{(\text{number of waves} \times \text{constant phase difference between two consecutive waves})}{2}}{\sin\left(\frac{\text{constant phase difference between two consecutive waves}}{2}\right)} \quad (12)$$

or

$$R' = \frac{n\,a \sin\alpha}{\alpha} = \frac{A \sin\alpha}{\alpha}, \quad (13)$$

Where n are the number of Huygen's secondary waves originated from each slit. A is the amplitude of each secondary wave.

And $\alpha = \frac{\pi}{\lambda} d \sin\theta$.

Where d is the width of each slit. θ is the angle in which direction the diffraction effect is to be analyzed from the horizontal.

Now we take two waves each of amplitude R' reaching at P and the phase difference between them is given by;

$$\phi = \frac{2\pi}{\lambda}(e + d)\sin\theta = 2\beta \quad (14)$$

As we can find the resultant of n number of waves, each having amplitude a and phase difference given by

$$R' = \frac{a \sin\frac{n\delta}{2}}{\sin\frac{\delta}{2}}.$$

Hence, the resultant of two waves each having amplitude R', given by (12) or (13) and phase difference 2β, given by (14), can be written as

$$R = \frac{R' \sin\frac{2 \times 2\beta}{2}}{\sin\frac{2\beta}{2}} \quad \text{or}$$

$$R = \frac{A \sin\alpha}{\alpha} \frac{\sin 2\beta}{\sin\beta}$$

So, the intensity due to two slits is given by

$$I = R^2 = A^2 \frac{\sin^2\alpha}{\alpha^2} \frac{\sin^2 2\beta}{\sin^2\beta} \quad \text{or}$$

$$I = A^2 \frac{\sin^2\alpha}{\alpha^2} \frac{4\sin^2\beta \cos^2\beta}{\sin^2\beta} \quad \text{or}$$

$$I = 4A^2 \frac{\sin^2\alpha}{\alpha^2} \cos^2\beta \quad (15)$$

This is the same expression as in the case of YDSE

$I = 4 I_0 \cos^2 \frac{(2\beta)}{2}$, where $I_0 = A^2 \frac{\sin^2\alpha}{\alpha^2}$, is the intensity of each slit.

$$I = 4 I_0 \cos^2 \frac{(\phi)}{2} \quad (16)$$

The equation (15) contains two terms, one i.e., $A^2 \frac{\sin^2\alpha}{\alpha^2}$, is the intensity of single slit diffraction and $4\cos^2\beta$, is intensity because of two slit interference. So, the maxima and minima of intensity due to diffraction through double slits, will be explained by explaining the single slit diffraction term i.e., $A^2 \frac{\sin^2\alpha}{\alpha^2}$, and the double slit interference term i.e., $4\cos^2\beta$.

(A) Maxima and Minima due to individual single slit

In equation (15), the term, $A^2 \frac{\sin^2\alpha}{\alpha^2}$, is due to each single slit diffraction. The maxima and minima due to single slit has been discussed in section 4.3.4. There, we have seen that;

(i) Condition for minimum is $d\sin\theta = \pm n\lambda$, and the minima will be obtained in the direction; $\theta = \pm\sin^{-1}(\frac{n\lambda}{d})$

(ii) The condition for maxima is $\alpha = \tan\alpha$, and the approximate values of α will be;

$\alpha = \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$ etc.

(B) Maxima and Minima due to two slit interference term i.e., $\cos^2\beta$ is;

(i) For Minima, $\cos^2\beta = 0$

Or $B = \pm(2n+1)\frac{\pi}{2}$

Or $$(e+d)\sin\theta = \pm(2n+1)\frac{\lambda}{2} \qquad (17)$$

(ii) For maxima, $\cos^2\beta = 1$

Or $\beta = \pm n\pi$

Or $$(e+d)\sin\theta = \pm n\lambda \qquad (18)$$

And the intensity curve is shown in Fig. 4.6.

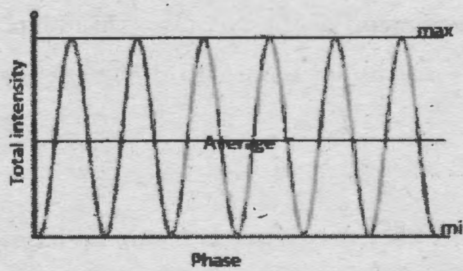


Fig. 4.6 Intensity variation curve due to double slit interference term, $\cos^2\beta$

Now according to equation (15) the resultant curve of Fraunhofer diffraction at two slits should be the superposition of two curves, Fig. 4.5 and Fig. 4.6 and will be drawn as Fig. 4.7.

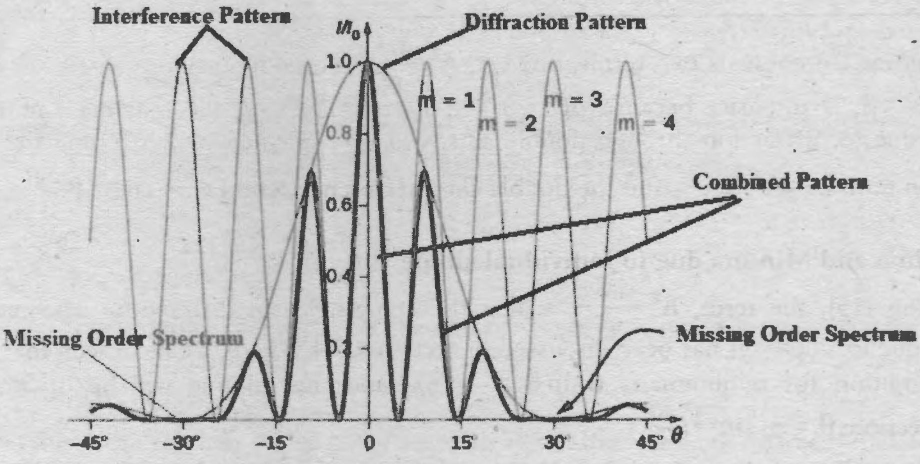


Fig. 4.7 Overlapping of Single slit Diffraction and two slit interference Pattern

4.4.1 Number of Interference Maxima Inside Central Maxima of Single Slit

The angular half width of a central maximum in a single slit is given by;

$$\theta = \pm \sin^{-1}\left(\frac{\lambda}{d}\right)$$

or $$\theta = \frac{\lambda}{d}, \text{ if } \theta \text{ is small.}$$

So, the width will be

or $$2\theta = 2\,\frac{\lambda}{d},$$

In case of YDSE the position of n^{th} maximum is

$$y_n = \frac{nD\lambda}{2d}, \text{ or}$$

Here it should be noted that, 2d, the distance between two slit is $(e + d)$, d is the width individual slit and e is the opacity between them, so

$$y_n = \frac{nD\lambda}{(e+d)},$$

2θ can also be written as,

$$2\theta = \frac{2y_n}{D} = 2\,\frac{\lambda}{d}$$

Now putting the value of $\frac{y_n}{D}$, we get;

$$\frac{2n\lambda}{(e+d)} = 2\,\frac{\lambda}{d}$$

So, the total number of Interference Maxima $= 2\,\frac{(e+d)}{d}$ (19)

As 'n' is the number of maxima of each side of the central point, so 2n will be the number of interference maxima in the central maximum of a single slit and depends on the ratio of the opacity to the slit of the width.

Special note

From above discussion, we can say that YDSE is nothing but the superposition of two diffraction patterns from two adjacent single slits. It means as soon as we want to treat the interference through two slits (i.e., YDSE), actually we should treat the diffraction from each single slit and then superimposes the intensity patterns of the diffraction of two single slits. Here the point that comes into mind is that the intensity pattern obtained by YDSE and diffraction at double slit should be same. But they seem to be different because while performing Young's Double Slit Experiment, the intensity variation due to diffraction, at individual slit was not taken into consideration. Here, also it can be seen from equation (15), that if the term $A^2 \frac{\sin^2\alpha}{\alpha^2}$ is considered to be constant, I_0 , i.e., it does not vary along any direction, then the curve given in Fig. 4.7 will be similar to that of the intensity curve of Young's Double Slit Experiment.

4.5 Fraunhofer Diffraction at Circular Aperture

Fraunhofer diffraction through circular aperture was first analyzed by Airy in 1835. After doing comprehensive analysis, he found that it can be explained using the theory of single slit diffraction with some correction factor in finding the diameter of the diffraction rings obtained on the screen. These circular rings are called Airy's rings. A circular slit can be considered to be made by rotating a single slit about its centre of gravity. Now let AB is a circular slit of diameter d. Let S is a source of light and P is the screen. The lenses are used to obtain the condition for Fraunhofer diffraction.

As each diametric line AB can be considered as a single slit and the conditions for minima and maxima are given by;

$d\sin\theta = \pm\, n\lambda$ (Condition for minima) (20)

$d\sin\theta = \pm(2n+1)\frac{\lambda}{2}$ (Condition for maxima) (21)

The obtained intensity pattern i.e., alternate bright and dark concentric rings, can be considered as a result of rotation of the single slit pattern, about the central maximum as in Fig. 4.8. Now using the equations (20), (21) and Fig.4.8, we can find the diameter of centre bright ring.

For the large distance, the angular width of central maxima due to single slit is written as;

$d\sin\theta = \pm\lambda$

We have,

$\theta = \pm\sin^{-1}(\frac{\lambda}{d})$, thus the central maximum extends from $+\sin^{-1}(\frac{\lambda}{d})$ to $-\sin^{-1}(\frac{\lambda}{d})$.

In other words, 2θ, is called the angular width of the central maximum i.e.

Angular width of central maximum is $2\theta = 2[\pm\sin^{-1}\left(\frac{\lambda}{d}\right)]$

If the screen is at a large distance, then, the width of central maxima will be;

$$\theta = \left(\frac{\lambda}{d}\right) \tag{22}$$

Now if the lens is very near to the slit and the screen is the focus of the lens then the linear width of the bright fringe can be calculated as;

$$\theta = \left(\frac{x}{f}\right) \tag{23}$$

From (22) and (23) we have, the linear width of central maxima of the Fraunhofer diffraction through single slit as;

$$x = \left(\frac{f\lambda}{d}\right) \tag{24}$$

If the slit is in a circular aperture form, then the value of x i.e., the diameter of the central bright ring called Airy's ring, is modified with a correction factor called Airy's correction, as;

$$x = 1.22\ \frac{f\lambda}{d} \tag{25}$$

Where 1.22 is Airy's correction, introduced when a formula for a rectangular slit is used for circular slit.

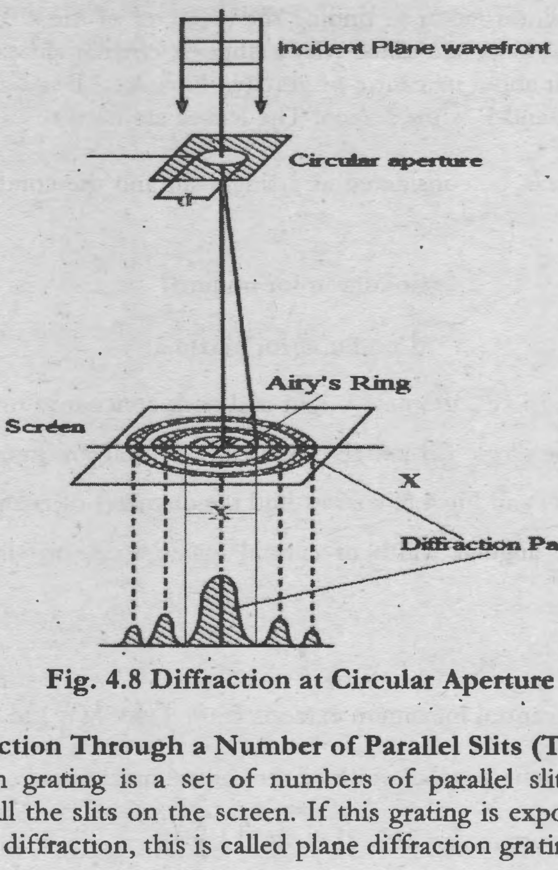


Fig. 4.8 Diffraction at Circular Aperture

4.6 Fraunhofer Diffraction Through a Number of Parallel Slits (Theory of Grating)

Grating: A diffraction grating is a set of numbers of parallel slits, showing the collective diffraction pattern of all the slits on the screen. If this grating is exposed to a plane wave front to produce Fraunhofer diffraction, this is called plane diffraction grating.

Let there are N numbers of parallel slits which are exposed to a plane wave front to analyze Fraunhofer type of diffraction. The grating can be made by drawing equidistant transparent and opaque lines on a glass plate. For diffraction to takes place though this grating, the width of transparency and the opacity should be of the order of wavelength of light. Although, it is very difficult to make such type of arrangement but with the help of laser such rulings can be drawn on a glass plate coated with a translucent film. Let d be the width of transparency and e is the width of opacity of the grating. This grating is illuminated with a parallel wave front and the pattern is obtained on the screen, as shown in the Fig. 4.9. Now let us find out the conditions for intensity variation on the screen.

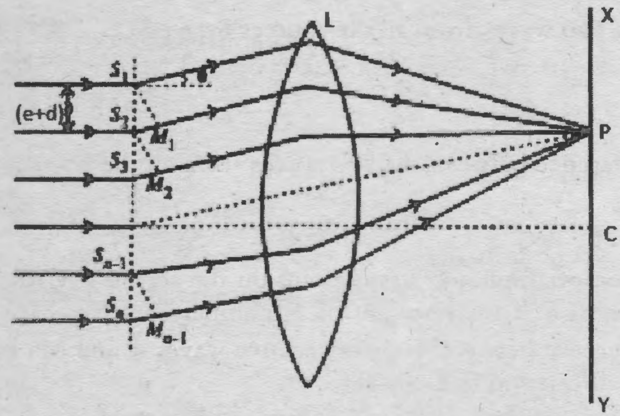


Fig. 4.9 Fraunhofer Diffraction through a number of parallel slits (Grating)

4.6.1 Analytical treatment of Fraunhofer Diffraction Through a Grating

Let a wave front of monochromatic light is incident on a grating. The parallelism is obtained with the help of lenses where needed. As the plane wave front reaches the grating i.e., a combination of slits, each slit shows its diffraction pattern on the screen. Obviously, the combined effect of all the slits is seen on the screen called as diffraction pattern which is obtained on the screen due to a grating (N number of slits). We have already worked out in the previous section the resultant amplitude at a point on the screen, considering Fraunhofer diffraction through single slit. The result can be applied to find the amplitude due to diffraction at each individual slit on the screen. Now, if we want to find out the amplitude of light at a point in the pattern obtained on the screen, we simply have to add the amplitudes of all the individual slits at that point. All the slits are exactly similar here and the separation between the consecutive slits is also very small. Therefore, we may consider that the amplitude of each individual slit at any point on the screen is equal. (e + d) will be the separation between two corresponding points in the two consecutive slits. This distance is called grating constant. Let R' be the amplitude due to an individual slit at a point P on the screen, which is already calculated for the case of diffraction at single slit, as;

$$R' = \frac{a \sin \frac{n\delta}{2}}{\sin \frac{\delta}{2}} \qquad (26)$$

Or, $R' = \frac{A \sin \alpha}{\alpha}$ (27)

Where a is the amplitude of each of the Huygen's secondary sources, generated at the single slit, n is the number of Huygen's sources on each slit. The phase difference between two consecutive waves is given by:

$\delta = \frac{2\pi}{\lambda}\frac{d}{n}\sin\theta$,

Where θ is the direction in which the diffraction effect is to be measured.

As there are N numbers of slits, each sending a wave of amplitude R' at a point on the screen with equal phase difference between two the consecutive sources, thus, the phase difference can be calculated as follows:

Path difference between two waves from all the sources $S_1, S_2, S_3, \ldots\ldots\ldots\ldots S_N$ will be given as:

$S_2M_1 = S_3M_2 = S_4M_3 \ldots\ldots\ldots\ldots\ldots\ldots = S_NM_{N-1} = (e+d)\sin\theta$.

Or

The common phase difference between the two waves from all the sources $S_1, S_2, S_3, \ldots\ldots\ldots S_N$ will be given as:

$\phi = \frac{2\pi}{\lambda}(e+d)\sin\theta = 2\beta$ (let) (28)

Now to find the equation of amplitude at any point on the screen due to a grating (N slits), we can treat the problem as to find the resultant of N number of waves each having amplitude R' and constant phase difference between two consecutive waves ϕ and can be find using equation (7) of the section under diffraction at single slit.

$$R = \frac{R' \sin \frac{n\phi}{2}}{\sin \frac{\phi}{2}}$$

$$R = \frac{R' \sin N\beta}{\sin\beta}$$

So, $R = \frac{A \sin\alpha}{\alpha}\frac{\sin N\beta}{\sin\beta}$ (29)

This clearly shows that the amplitude at any point on the screen due to N numbers of slits, comprises two terms, one term i.e., $\frac{A\sin\alpha}{\alpha}$ is the amplitude due to diffraction at a single slit and the term $\frac{\sin N\beta}{\sin\beta}$ may be taken as the combined effect of N identical slits. Obviously, the intensity due to grating is given as:

$I = R^2 = A^2\frac{\sin^2\alpha}{\alpha^2}\frac{\sin^2 N\beta}{\sin^2\beta}$ (30)

Now we have to discuss the conditions of maximum and minima formed at the screen. As the directions for maxima and minima formed due to $A^2\frac{\sin^2\alpha}{\alpha^2}$ is already been discussed in the single slit diffraction, so we only have to discuss the contribution of the term $\frac{\sin^2 N\beta}{\sin^2\beta}$ on the pattern obtained due to single slit.

(i) If the term $\frac{\sin^2 N\beta}{\sin^2\beta}$ gives its maximum value, then it is possible only when

$\sin\beta = 0$ i.e.

$\beta = \pm n\pi$ (n = 0,1,2,3......etc.) these values of β makes $\sin N\beta = 0$ and $\frac{\sin N\beta}{\sin\beta}$ becomes indeterminate. So, the value of $\frac{\sin N\beta}{\sin\beta}$ will be calculated by L'Hospital's rule as follows;

$$\lim_{\beta\to n\pi} \frac{\sin N\beta}{\sin\beta} = \lim_{\beta\to n\pi} \frac{\frac{d}{d\beta}(\sin N\beta)}{\frac{d}{d\beta}\sin\beta}$$

$$= \lim_{\beta\to n\pi} \frac{N\cos N\beta}{\cos\beta} = N$$

Putting $\lim_{\beta\to n\pi} \frac{\sin N\beta}{\sin\beta} = N$ in equation (16),

The intensity at any point P will be:

$$I = R^2 = A^2 \frac{\sin^2\alpha}{\alpha^2} N^2 \qquad (31)$$

The equation (31) shows that in the directions decided by $\beta = \pm n\pi$ (n= 0,1,2,3......etc.), the intensities of single slit diffraction patterns will be increased by N^2 times whatever it is, whether it is maximum or minimum in that particular direction.
Let us find out the directions for maxima.

(i) Principle Maxima
As for principle maxima $\sin\beta = 0$, i.e.

$$\beta = \pm n\pi \qquad (n = 0,1,2,3......etc.),$$

$$\beta = \frac{\pi}{\lambda}(e + d)\sin\theta = \pm n\pi$$

$$(e + d)\sin\theta = \pm n\lambda \qquad (32)$$

Or

$$\theta = \pm \sin^{-1}\left(\frac{n\lambda}{e + d}\right) \qquad (33)$$

For n = 0, $\theta = 0°$. This is the direction where the zero-order principle maximum of a single silt is formed. It is clear that the intensity of zero order principle maximum will be increased by a factor of N^2 where N is the number of slits on the grating. It is a matter of discussion that, it might not be true that the intensity of all the maxima will be increased by N^2 times. It is true only for, where the directions of principle maxima due to single slit are same as that of direction in which the term $\frac{\sin N\beta}{\sin\beta}$ gives its maximum value as N^2. But it is sure that the intensity (whatever it is) of single slit diffraction pattern will be increased by N^2 times in the direction $\theta = \pm\sin^{-1}\left(\frac{n\lambda}{e+d}\right)$, as this is not the direction of maxima due to single slit.

(ii) Minima
From equation (30), it is clear that intensity will be minimum, if $\sin N\beta = 0$, but $\sin\beta \neq 0$, then to find the points of maxima;
$\sin N\beta = 0$

$\Rightarrow \qquad N\beta = \pm m\pi$

Or $\qquad N\frac{\pi}{\lambda}(e + d)\sin\theta = \pm n\pi$

Or $\qquad N(e + d)\sin\theta = \pm m\lambda \qquad (34)$

Or $\theta = \pm\sin^{-1}\left(\frac{m\lambda}{N(e+d)}\right)$ are the directions in which $\frac{\sin N\beta}{\sin\beta}$ will be zero, if the values of m are except 0, N, 2N, ...etc.

From the above discussion, it is clear that one of the effects of increasing the number of slits in place of single slit is that there will be an increase in the intensity of the intensity pattern of a single slit by a factor of N^2. There may be some directions in which the maxima due to single slit falls upon minima due to the N slits. These maxima hence will not be clearly visible in the pattern. In other words, we can say that these maxima due to single slit are missing in the pattern. This may be due to the presence of another effect due to diffraction with N slits. This is called missing order pattern, which will be discussed later separately.

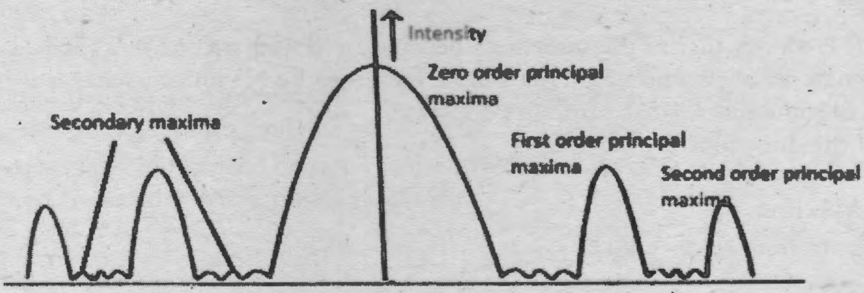


Fig 4.10 Intensity Distribution curve for plane diffraction grating

(iii) Concept of secondary maxima

It is clear from equation (34) that m = 0, N, 2N, ...etc. gives the same conditions as that are obtained from equation (32), on putting n = 0,1,2.......etc. and it implies that m = 0 corresponds to zero order principle maximum and m = N corresponds to first order principle maximum, etc. it means, for all the values of m which are in between 0 and N, N and 2N, etc., there are N-1 equidistant minima between zero order and first order principle maxima. It is obvious that there should be N-2 other maxima between zero order and first order principle maxima. These maxima are generated due to N slits, called as secondary maxima. These are of practically no interest due to very low intensity in comparison to the intensity of the principle maxima. The intensity of these maxima can be determined as follows;

Differentiating equation (30) w.r.t β, we get:

$$\frac{dI}{d\beta} = A^2\frac{\sin^2\alpha}{\alpha^2}\,\frac{\sin N\beta}{\sin\beta}\left[\frac{N\cos N\beta\sin\beta - \sin N\beta\cos\beta}{\sin^2\beta}\right] = 0$$

Or $\; N\cos N\beta\sin\beta - \sin N\beta\cos\beta = 0$

Or $\qquad \tan N\beta = N\tan\beta \qquad (35)$

This is the condition for secondary maxima. The intensity of the maxima can be found by determining the value of the term $\frac{\sin^2 N\beta}{\sin^2 \beta}$ in context with $\tan N\beta = N \tan \beta$.

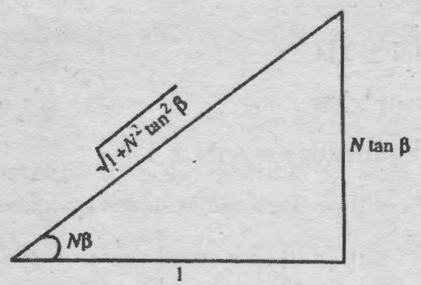


Fig. 4.11

Now, using Fig.4.11. We have

$$\sin N\beta = \frac{N \tan\beta}{\sqrt{(1+N^2\tan^2\beta)}}$$

So $$\frac{\sin^2 N\beta}{\sin^2\beta} = \frac{N^2\tan^2\beta}{\sin^2\beta((1+N^2\tan^2\beta)} = \frac{N^2}{1+\sin^2\beta\,(N^2-1)}$$

Thus, the intensity of secondary maxima is given by:

$$I^s = A^2\frac{\sin^2\alpha}{\alpha^2}\,\frac{N^2}{1+\sin^2\beta\,(N^2-1)} \tag{36}$$

To make a comparison let us take the ratio of intensity of secondary to the intensity of primary maxima. i.e.

$$\frac{I^s}{I} = \frac{1}{1+\sin^2\beta\,(N^2-1)} \tag{37}$$

As N increases, the intensity of secondary maxima decreases. Thus, if N is very large, the secondary maxima are not visible in the spectrum. That's why in case of a grating, there is uniform darkness between the two primary maxima.

4.6.2 Other Merits of Large Number of Slits (N) in a Grating

The width of zero order maxima in a single slit pattern is very large, to the width of other maxima which are half of the width of the central maximum. But here we can see that the width of any maxima reduces to a spectral line which can help us in taking the measurement for the ultimate use of a grating.

4.6.3 Width of any Principle Maxima

As any maxima is in between two minima, so angular separation between two minima will give the angular width of the corresponding maxima.

Suppose that we have to find out the angular width of n^{th} maxima formed along the direction θ_n. Then the equation of n^{th} maxima is given as;

$$(e + d) \sin \theta_n = \pm n\lambda \tag{38}$$

Now as n^{th} maxima correspond to nN^{th} minima, so the equations for adjacent minima to the n^{th} maxima will be written as:

$$N(e + d) [\sin(\theta_n \pm d\theta_n)] = (nN \pm 1)\lambda$$

or

$$N(e + d)[\sin \theta_n \cos d\theta_n \pm \cos \theta_n \sin d\theta_n) = (nN \pm 1)\lambda \tag{39}$$

Here, $+ d\theta_n$ to $- d\theta_n$ i.e., $2\, d\theta_n$ will be the angular width of n^{th} maxima.

$$\sin d\theta_n = d\theta_n \text{ and } \cos d\theta_n = 1, \text{ for small value of } d\theta_n.$$

So, equation (39) takes the form

$$N(e + d) \sin \theta_n \pm N(e + d) \cos \theta_n . d\theta_n = nN\lambda \pm \lambda \tag{40}$$

Multiplying equation (38) by N and subtracting from (40) we get;

$$N(e + d) \cos \theta_n . d\theta_n = \lambda$$

Or $$d\theta_n = \frac{\lambda}{N\,(e+d) \cos \theta_n} \tag{41}$$

Or $$2d\theta_n = \frac{\lambda}{N\,(e+d) \cos \theta_n},$$

will be the angular width of any principle maximum. This is inversely proportional to N. Most of the gratings have N, very large, say 15000. So, the width will be squeezed to a spectral line. This is one of the major advantages of choosing very large number of slits on the grating. Fine spectral line will be very useful in taking measurements for finding angle of diffraction for any order of the diffraction pattern with the help of spectrometer. Otherwise, if we choose a large width of any order of diffraction pattern, it will be very difficult to find accurate angle of diffraction for a particular order. Also, if two wavelengths are very close to each other, it will be very difficult to find angle of diffraction from them separately. This phenomenon comes under the resolving power of a grating discussed in the next topic.

4.7 Overlapping of Spectral Lines

It is seen that the grating exhibits a large range of wavelengths lying between 4000 Å to 7800 Å i.e., the visible spectrum. It is quite a possibility that overlapping of spectral lines might occur i.e., shorter wavelength spectral line of higher order may overlap with a longer wavelength spectral line of lower order. This is due to the fact that the equation of n^{th} principal maxima for wavelength λ in a diffraction grating given as: $(e + d) \sin \theta = n \lambda$ may be simultaneously satisfied by large values of λ with appropriate values of n.

If the angle of diffraction θ is same for say the spectral line of wavelength λ_1 for n_1 order, for the spectral line of wavelength λ_2 for n_2 order and so on…..then we have:

$$(e + d) \sin \theta = n_1 \lambda_1 = n_2 \lambda_2 = n_3 \lambda_3 \ldots\ldots\text{etc.} \tag{42}$$

This implies that the spectral line of wavelength λ_1 for n_1 order coincides with the spectral line of wavelength λ_2 for n_2 order and wavelength λ_3 for n_3 order and so on. These three lines occupy the same position in the spectrum then. However, no overlapping is observed for the visible light extending from 4000 Å to 7800 Å for the first and second orders. But the overlapping is more pronounced for higher orders.

e.g., The red light of wavelength 7000 Å in third order, the green light of wavelength 5250 Å in fourth order and violet light of wavelength 4200 Å in fifth order coincide as:

$$(e+d)\sin\theta = 3 \times 7000 \times 10^{-8} = 4 \times 5250 \times 10^{-8} = 5 \times 4200 \times 10^{-8}$$

When overlapping of spectra occurs, it is not possible to study any individual spectrum. This difficulty can be overcome by the use of suitable filters to absorb the undesired wavelengths of incident light which may overlap with the spectral lines under investigation.

4.8 Highest Wavelength to Observe Diffraction Through Single Slit

Condition for minima through a single slit is:

$$d \sin\theta_n = \pm n\lambda$$

$$\sin\theta_n = \pm\frac{n\lambda}{d}$$

The highest value of λ may be equal to the slit width, only for n = 1. In that condition, $\theta = \pm\frac{\pi}{2}$, it means first minimum will be obtained on either side of the central maximum. Practically, whole of the screen will be illuminated with central maximum and it is to be believed that no diffraction is taking place. It means visible diffraction pattern is obtained only for $\lambda \leq$ slit width.

4.9 Highest Wavelength to Observe Diffraction Through a Grating

$$(e+d)\sin\theta_n = \pm n\lambda$$

$$\sin\theta_n = \pm\frac{n\lambda}{(e+d)}$$

This shows the wavelength $\lambda \leq$ grating constant.

4.10 Maximum Number of Orders Observable With a Diffraction Grating

The condition for maxima of diffraction through a grating is:

$$(e+d)\sin\theta_n = \pm n\lambda$$

Or the order of the spectrum, $n = \frac{(e+d)\sin\theta_n}{\lambda}$

For the maximum possible value of diffraction θ is 90° for which we have $\sin\theta = 1$. Therefore, the maximum number of possible orders is given by:

$$n_{max} = \frac{(e+d)}{\lambda}$$

The ratio of this grating element and the wavelength of light is dependent on n_{max}

e.g., when grating element $(e+d) < 2\lambda$, for normal incidence, we have:

$$n_{max} < \frac{2\lambda}{\lambda} < 2$$

This implies that for normal incidence when $(e+d) < 2\lambda$, only first order is visible.

4.11 Spectrum Obtained with a Grating

If a grating is exposed with a visible band of wavelength, a coloured diffraction pattern will be seen on the screen on either side of white central maximum as shown in Fig.4.12. Obviously, the pattern will consist of lowest wavelength towards the central maximum. This phenomenon is widely being used for measurement of any wavelength out of a visible band of wavelengths. The other characteristics of the spectrum includes:

(i) The spectral lines are sharp and bright.
(ii) On either side of the central maximum in the spectrum we have colours in the order from violet to red.
(iii) On either side of the central maximum, a number of spectra of different orders are situated symmetrically.
(iv) The condition for maxima of diffraction through a grating is $(e + d)\sin\theta_n = \pm n\lambda$. thus, for central maxima $n = 0$ and so $\theta = 0°$ for all the wavelengths. This implies that the central maxima for all the wavelengths coincide, forming its image same as that of the source of light.
(v) Most of the transmitted light is concentrated on the central maxima and the rest is distributed in the spectra of other orders. Thus, the spectra formed by grating are less intense as the intensity goes on decreasing as order becomes higher and higher.

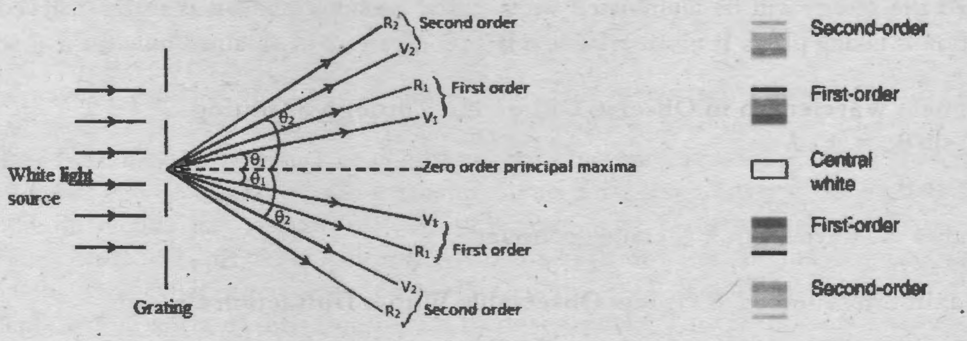


Fig.4.12. Grating spectrum

4.12 Missing Order of Spectrum

We know that the spectrum obtained through a grating is the superposition of diffraction patterns of a number of single slits and we have separate conditions for maxima and minima for superposition of a number of waves from Huygens, secondary sources (Single slit Diffraction) and superposition of a numbers of waves from N identical sources (Diffraction through a grating). Now this is quite a possibility that n^{th} minima due to single slit fall on the m^{th} maxima due to N slits. In this situation, the maxima may be missing from the spectrum. This phenomenon is called missing order spectrum. The condition for the same is obtained as follows:
The condition of minima for single slit is:

$$d\sin\theta_n = \pm n\lambda \qquad (43)$$

and the condition for maxima for N slits as:

$$(e + d) \sin \theta_n = \pm m\lambda \qquad (44)$$

Divide (38) by (37):

Then $\frac{e+d}{d} = \frac{m}{n}$ (45)

Or $m = n \left(\frac{e+d}{d}\right)^{th}$ order will be missing in the spectrum obtained from the grating.

The ratio of e : d will give the exact missing orders.

For example:

(i) If e = 2d, then 3rd, 6th, 9th order of maxima of grating will not be clearly visible and we cannot do measurements on these orders for the calculation of measurement of wavelength etc.

(ii) If e = d, then 2nd, 4th, 6th, 8th...order of maxima will be absent.

4.13 Applications of Phenomena of Diffraction

There are a lot of practical applications of diffraction of electromagnetic waves, such as X-ray diffraction spectroscopy, grating spectroscopy, measurement of slit width with laser, resolution of two close objects (slit/obstacle) or two close wavelengths, etc. Here we will discuss a method of measurement of wavelength of light with a grating.

4.14 Measurement of Wavelength with Grating

An appropriate grating placed on the spectrometer, is illuminated with the light of which wavelength is to be measured. Diffraction pattern by the grating is seen through telescope shown in Fig. 4.13.

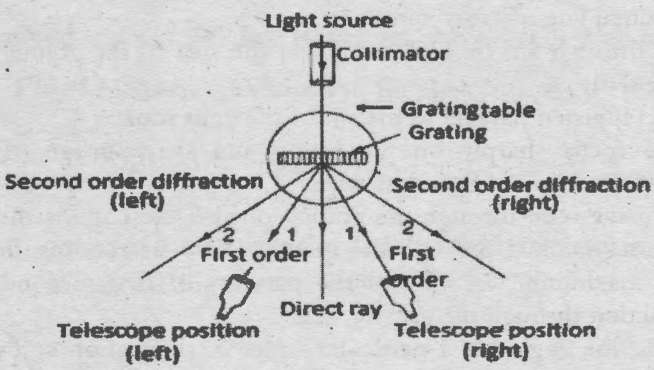


Fig.4.13. Experimental set up for measurement of wavelength using grating

Using the following equation for the maxima for grating, we can determine the desired wavelength. The equation for maxima for a grating is:

$$(e + d) \sin \theta_n = \pm n\lambda$$

Finding grating constant, angle of diffraction θ_n for a particular order, the wavelength λ can be measured.

(a) Measurement of Grating constant (e + d)

The grating constant is the distance between the two consecutive slits. Now if N are the number of slits and L is the total width of the grating. Then $\frac{L}{N}$ will be the distance between two consecutive slits i.e. grating constant

$$\text{Grating constant} = \frac{\text{Width of the grating}}{\text{Total No. of the slits}}$$

$$(e + d) = \frac{L}{N}\text{units}$$

(b) Measurement of Angle of Diffraction

To find the accurate angle of diffraction for a particular order, the following adjustment on the spectrometer is necessary to be made.

(i). The cross wire of the eyepiece of the telescope should be in focus of the eyepiece.
(ii) The telescope and the collimator are adjusted for infinity.
The grating is adjusted on the table of the spectrometer in such a way that the collimator and the telescope are exactly perpendicular to the grating so the light may fall on the grating so as to avoid oblique incidence. This is done using Schuster's method as follows:
The collimator and the telescope are so arranged that the direct image of the source (slit) falls on the vertical cross wire and the position through the telescope is noted.
The axis of the telescope is now kept perpendicular to that of collimator by rotating the telescope through 90^0.
The grating is now placed on the spectrometer table and the table is rotated for the position of the grating such that the reflected image of the slit is seen in the telescope. Using the levelling screws, the image of the slit is aligned vertically with the cross wire and also equally half above and below the horizontal line of cross wire.
The table is rotated through 45^0 or 135^0, such that the slits of the grating facing the light from the source perpendicularly.
The slits (rulings) are adjusted parallel to the slit of the light source.
Using the telescope focus sharply on the bright and sharp image of the source (slit) by coinciding with the vertical line of the cross wire.
Actually the direct image seen through the grating on the telescope is the central maximum of the diffraction pattern obtained through the grating. Now by rotating the telescope on either side of this central maximum, we will see the patterns of first, second, third etc. orders of maxima of the diffraction through the grating.
We can now measure the angle for a particular order of diffraction and note down the order. Putting the values of grating constant (e+d), angle of diffraction θ_n and the order of diffraction, wavelength λ can be measured.

4.15 Resolving Power of Optical Instrument

The ability of an optical instrument to resolve two close objects/wavelengths which are unresolved by the naked eye is called the resolving power of that optical instrument e.g. telescope, grating, prism etc.

Rayleigh Criterion

According to Rayleigh criterion "two close objects or two close wavelengths are said to be just resolved if the central maxima of one object/wavelength coincide with the first minima of other object/wavelength and vice versa. This is called Rayleigh criterion for minimum resolution Fig. 4.14 (B).

Let us consider two close wavelengths, λ and $\lambda + d\lambda$. If $d\lambda$ is quite large, the maxima of two wavelengths will be well resolved Fig. 4.14 (C). In this condition, the two wavelengths are well resolved. But if the value of $d\lambda$ is such that the maxima due to one wavelength falls upon the first minima of second wavelength, then this condition is said to satisfy Rayleigh criterion for minimum resolution. The two wavelengths are said to be in just resolved condition Fig. 4.14(B). But if the value of $d\lambda$ is below this value, the two wavelengths can not be resolved by any means Fig. 4.14(A).

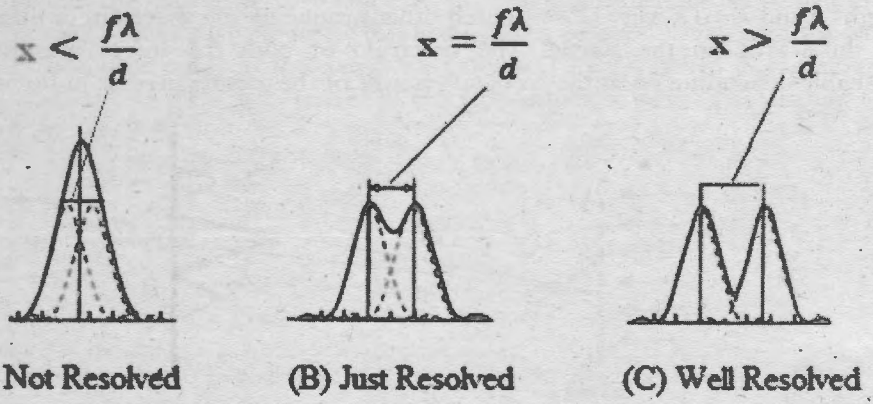


Fig. 4.14 Rayleigh Criterion for Resolution

We know the intensity of any maxima is given by:

$$I = R^2 = I_{max}\frac{\sin^2\alpha}{\alpha^2}$$

For first minima, $\alpha = \pi$, thus the intensity at the middle of the two maxima will be given at $\alpha = \frac{\pi}{2}$, and the intensity at that point due to each object will be

$$I_1 = I_2 R^2 = I_{max}\frac{\sin^2\frac{\pi}{2}}{\frac{\pi^2}{4}} = \frac{4}{\pi^2}I_{max}$$

The intensity in the middle of two maxima can be given as:

$$I_1+I_2=\frac{8}{\pi^2}I_{max}$$

In other words, if the intensity at the dip in diffraction pattern of two wavelengths is exactly equal to $\frac{8}{\pi^2}$ (i.e. 81%) of the intensity at the central maxima of either wavelengths, then the two wavelengths are said to be just resolved. If, the intensity is less than $0.81\, I_{max}$, the two wavelengths are well resolved. But if the intensity is greater than $0.81\, I_{max}$, then the two wavelengths cannot be resolved. This is one of the best methods to check whether the two wavelengths are resolved or not.

4.16 Resolving Power of a Grating

Resolving power of any optical instrument is its ability to resolve two close objects which could not be distinguished with naked eye. Its measurement is different for different optical instruments. Generally, it is inversely proportional to the minimum angular separation between the two objects that is made with eye. This angle further can be reduced with the help of telescope or microscope called the resolving power of these instruments. But in case of grating, it is the ability to resolve two close wavelengths. The criterion for minimum resolution for two wavelengths is given by Lord Rayleigh. According to which "two close wavelengths are said to be just resolved, if central maxima in the diffraction pattern of wavelength through single slit, coincides with the first minima of the diffraction pattern of other wavelength through the same slit". Using this criterion, the resolving power of a grating can be found as follows:

Let a grating of appropriate dimension is illuminated with a beam of light, consisting two wavelengths λ and $\lambda + d\lambda$, very close to each other producing the spectrum of diffraction pattern of both the waves on the screen. The spectrum of both the waves may or may not be distinguishable depending upon the resolving power of the grating used as in Fig. 4.15.

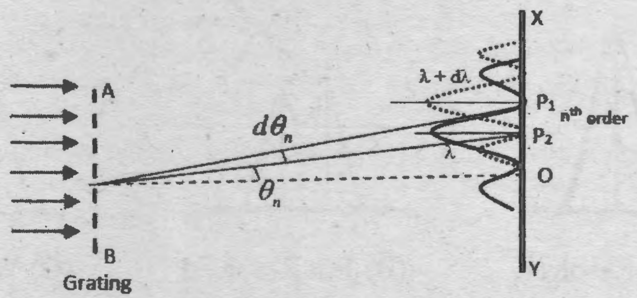


Fig. 4.15. Resolving power of a grating

If N be the number of slits in the grating and (e + d) is the grating constant/element. Writing the equations for maxima and minima for both the wavelengths, we have:

The condition for n^{th} maximum is for $\lambda + d\lambda$

$$(e + d)\sin\theta_n = \pm n(\lambda + d\lambda) \qquad (46)$$

To satisfy Rayleigh criterion if first minimum, adjacent to n^{th} maxima for wavelength λ, falls on the n^{th} maxima of $\lambda + d\lambda$, then the condition for first minimum adjacent to n^{th} maximum for wavelength $\lambda + d\lambda$ should be:

$$N(e + d)\sin\theta_n = (nN \pm 1)\lambda \qquad (47)$$

The direction θ_n should be same for n^{th} maximum for $\lambda + d\lambda$ and first minimum adjacent to n^{th} maxima for wavelength λ to satisfy Rayleigh criterion.

Multiplying equation (46) by N and subtracting (47) from it, we get;

$$[N(e + d)\sin\theta_n = \pm nN(\lambda + d\lambda)] - [N(e + d)\sin\theta_n = (nN \pm 1)\lambda]$$

$$nNd\lambda = \lambda$$

or $$\frac{\lambda}{d\lambda} = nN \tag{48}$$

$\frac{\lambda}{d\lambda}$ is called resolving power of a grating.

The equation (48) shows that more the numbers of slits on a grating, more is the resolving power of the grating and as we move towards the higher orders of spectrum, the patterns of two close wavelengths are more and more distinguishable. It also shows that resolving power is independent of grating constant.

4.17 Dispersive Power of Grating

It is clear from the expressions of the conditions of maxima and minima in the phenomena of diffraction that higher the wavelength, greater is the angle of diffraction. Obviously, the condition for diffraction to take place must be fulfilled i.e., the greatest wavelength for which we can see diffraction should be equal to the size of the object. This is verified in the next article. Now we have to find out the dispersive power of a grating which is defined as change of angle of diffraction with the wavelength.

The condition for maxima of diffraction through a grating is:

$$(e + d)\sin\theta_n = \pm n\lambda \tag{49}$$

Differentiating w. r. t λ, we get;

$$(e + d)\cos\theta_n \frac{d\theta}{d\lambda} = n \tag{50}$$

Or $$\frac{d\theta}{d\lambda} = \frac{n}{(e+d)\cos\theta_n} \tag{51}$$

This shows that more dispersion is visible in higher orders. Smaller the grating constant greater the dispersive power of the grating.

Difference between the resolving power and dispersive power of grating

The resolving power is given as:

$$\frac{\lambda}{d\lambda} = nN \tag{52}$$

And the dispersive power is given as:

$$\frac{d\theta}{d\lambda} = \frac{n}{(e+d)\cos\theta_n} = nN' \tag{53}$$

Where n is the order of the spectrum, N is the total number of lines on the grating, N′ is the total number of lines per cm on the grating surface, θ defines the direction of the n^{th} principal maxima corresponding to a wavelength λ. As we can see, the resolving power increases with increase in the number of slits (N) in the grating i.e., the width of the grating surface and the dispersive power on the other hand increases with increase in the number of lines per cm. If, for the two gratings N′ is same, then the dispersive power is same for both the cases but the grating having larger width will have larger resolving power giving rise to higher resolution of spectral lines. Larger the width of the grating surface, sharper and narrower the spectral lines be. The difference between the two can be clearly understood by the given Fig.4.16. In the Fig, we can see the diffraction maxima of two wavelengths formed by two diffraction gratings having same grating element (e+d) but different widths of the ruled surface. The angular dispersive power is same in both the cases but the resolving power is more in the second case as the two maxima are sharp and a much smaller difference in the wavelengths can be detected.

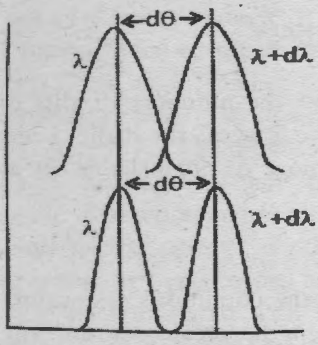


Fig.4.16. Difference between resolving power and dispersive power of grating

4.18 Merits of a Grating over a Single Slit

As the basic difference between the two is the number of slits so all the benefits of increasing the number of slits are the merits of grating over a single slit. Here are some of them:

The intensity of all the maxima are increased as $I \propto N^2$.

Angular width of all the maxima are reduced as $2d\theta_n \propto 1/N$. In case of very large N, all the maxima are squeezed in the form of spectral lines. So the measurement of angle of diffraction for a particular order of maximum will be very accurate.

Resolving power is increased as R.P $\propto$ N.

4.19 Similarities and Dissimilarities between Interference and Diffraction

Similarities

Both the phenomena occur due to the superposition of light waves.

Both shows the regions of varying intensity on the screen called fringes.

Dissimilarities

Interference occurs due to superposition of two waves from two separate identical sources whereas diffraction occurs due to the superposition of two or a number of waves originating from same wave front.

Intensity variation in the interference pattern is completely sinusoidal. The intensity of all successive maxima and minima in the interference pattern will be constant in whole of the pattern if and only if the two waves from two identical sources are not taken due to the resultant of diffraction occurring from individual single slits. Otherwise, the amplitude of each wave from a single slit will vary at the different points on the screen, according to the single slit diffraction phenomena. And the combined effect of diffraction pattern from two single slits will be obtained on the screen.

Intensity variation in the diffraction pattern is exponentially decaying of successive maxima.

*******Solved Examples*******

Based on diffraction through a single slit

Ex.1 Light of wavelength 5500 Å falls normally on a slit width 22.0×10^{-5} cm. Calculate the angular position of the first two minima on either side of the central maximum.

Sol. The angular position of minima in the diffraction pattern produced by a single slit are given by

$$d \sin\theta = n\lambda \quad \text{or} \quad \sin\theta = \frac{n\lambda}{d}$$

where n is the order of minima and d is the width of the slit.

Here, $d = 22.0 \times 10^{-5}$ cm. and $\lambda = 5500 \times 10^{-8}$ cm.

For first order $n = 1$,

$$\sin\theta_1 = \frac{\lambda}{d} = \frac{5500 \times 10^{-8}}{22.0 \times 10^{-5}} = 0.25$$

$$\therefore \quad \theta_1 = \sin^{-1}(0.25) = 14°29'.$$

For second order $n = 2$,

$$\therefore \quad \sin\theta_2 = \frac{2\lambda}{d} = \frac{2 \times 5500 \times 10^{-8}}{22.0 \times 10^{-5}} = \frac{1}{2}$$

$$\therefore \quad \theta_2 = 30°$$

Thus, the first two minima will occur at the angles 14° 29′ and 30° on either side of the central maximum.

Ex.2 Plane waves of $\lambda = 6.0 \times 10^{-5}$ cm fall normally on a straight slit of width 0.20 mm. Calculate the total angular width of the central maximum and also the linear width as observed on a screen placed 2 metres away.

Sol. The angular half width θ of the central maximum is given by

$$\sin\theta = \frac{\lambda}{d}$$

Here $\lambda = 6.0 \times 10^{-5}$ cm and $d = 0.20$ mm $= 0.02$ cm.

$$\sin\theta = \left(\frac{6.0 \times 10^{-5}}{0.2}\right) = (3 \times 10^{-3})$$

As $\sin\theta$ is very small, we may write

$$\theta = \sin\theta = 3 \times 10^{-3} \text{radians}$$

$\therefore$ The total angular width of the central maximum

$$= 2\theta$$
$$= 2 \times 3 \times 10^{-3}$$
$$= 6 \times 10^{-3} \text{ radians}$$

If y is the linear half width and d is the distance of the screen from the slit, we have

$$y = d\theta \text{ as } \theta \text{ is small}$$

Here $d = 2$ meters $= 200$ cm. and $\theta = 3 \times 10^{-3}$ radians.

$$\therefore \quad y = 200 \times 3 \times 10^{-3} \text{ cm.} = 0.6 \text{ cm.}$$

$\therefore$ The linear width of the central maximum on the screen

$$= 2y = 2 \times 0.6 = 1.2 \text{ cm}$$

Based on diffraction at two slits and circular aperture

Ex.3 In a diffraction phenomenon using double slit, calculate the (i) distance between the central maximum and the first minimum of the fringe envelope and (ii) the distance between any two consecutive double slit dark fringes.

Given: wavelength of light = 5000 **Å**, spacing between two slits = 0.10mm, screen to slits distance = 100cm

Sol. e = 0.02 mm $= 2\times10^{-5}$m, d = 0.1mm = 10^{-4}m, (e + d) = 1.2×10^{-4}m, $\lambda = 5000\text{Å} = 5\times10^{-7}$m, D = 100 cm = 1m

(i)The angular separation between the central maximum and the first minimum is,

$\sin\theta_1 = \theta_1 = \frac{\lambda}{2(e+d)}$ and $\theta_1 = \frac{x_1}{D}$

Or $\frac{x_1}{D} = \frac{\lambda}{2(e+d)}$ or $x_1 = \frac{\lambda D}{2(e+d)}$

$$x_1 = \frac{5\times10^{-7}\text{m}\times1\text{m}}{2(1.2\times10^{-4}\text{m})} = 2.08\times10^{-3}\text{m} = 2.088\text{ mm}$$

(ii)The angular separation between two consecutive dark fringes,

$$\sin\theta_1 - \sin\theta_2 = \theta_1 - \theta_2 = \theta = \frac{3\lambda}{2(e+d)} - \frac{\lambda}{2(e+d)}$$

$\theta = \frac{\lambda}{(e+d)}$ also $\theta_2 = \frac{x_2}{D}$

Or $x_2 = \frac{\lambda D}{(e+d)}$ or $x_2 = \frac{5\times10^{-7}\text{m}\times1\text{m}}{(1.2\times10^{-4}\text{m})} = 4.16\times10^{-3}\text{m} = 4.2\text{mm}$

Ex.4 In a double slit Fraunhoffer diffraction pattern, the screen is placed 170cm away from the slits. The width of the slits is 0.08mm and they are 0.4mm apart. Calculate the wavelength of light if the fringe width is 0.25cm. Also, find the missing order.

Sol. (i) The fringe width is given by:

$\beta = \frac{\lambda D}{(e+d)}$ or $\lambda = \frac{\beta(e+d)}{D}$

Given: e = 0.08mm = 0.008cm, d = 0.4mm = 0.04cm,, D = 170 cm, β =0.25cm

Thus, $\lambda = \frac{0.25\times(0.008+0.04)}{170}\text{cm} = 7.059\times10^{-5}\text{cm}$

Or $\lambda = 7059\text{ Å}$

(ii)The condition of missing order is

$$\frac{e+d}{e} = \frac{n}{m}$$

$\frac{0.008+0.04}{0.008} = \frac{n}{m}$ or n = 6m

Or n = 6, 12, 18,........

Hence 6th, 12th, 18th,................orders are missing.

Ex.5 In a double slit arrangement each slit has a width of 0.02mm and the distance between the slit is 0.10mm. Mercury blue light of wavelength 4358 **Å**, obtained using a filter, is incident normally on the double slit. Calculate (a) the fringe spacing and (b) the linear distance from the central maxima to first minimum if the pattern is formed on the screen 60cm away from the slits.

Sol. (a) Angular fringe width = = $\frac{\lambda}{(e+d)}$

Thus, linear fringe spacing Δy at a distance D = $\frac{\lambda D}{(e+d)}$

Or $\quad \Delta y = \frac{4358 \times 10^{-10} m \times 0.6m}{0.1 \times 10^{-3} m} = 2.6 \times 10^{-8} m$

(b)The angular position θ of the first minimum is given by:

$$\sin\theta \cong \theta = \frac{\lambda}{e} = \frac{4358 \times 10^{-10} m}{0.02 \times 10^{-3} m} \approx 0.022 \text{ radian}$$

Linear separation $\Delta y' = \theta D = 0.022 \times 0.6m = 0.013m$

Now, $\quad \frac{\Delta y'}{\Delta y} = \frac{0.013}{2.6 \times 10^{-3}} = 5$

So, there are 10 fringes in the central maximum of the diffraction envelope.

Ex.6 Find the angle subtended by the Airy disk in the Fraunhofer diffraction pattern of a circular aperture of diameter 0.20mm illuminated by a light of He-Ne laser (λ = 633nm)

Sol. Given: D = 0.20mm, λ = 633nm

$$\theta_{airy} = \frac{1.22\lambda}{D}$$

We get, $\quad \theta_{airy} = \frac{1.22 \times 633 \times 10^{-9}}{0.20 \times 10^{-3}} = 3861.3 \times 10^{-6} = 0.0038 \text{ degree}$

Based on diffraction grating

Ex.7 Light of wavelength 5000Å is incident normally on a plane transmission grating of width 3 cm. and 15000 lines. Find the angle of diffraction in first order.

Sol. The direction of principal maxima for normal incidence are given by

$$(e+d)\sin\theta = n\lambda$$

where (e+ d) is grating element and n is the order of maxima.

Here, $\quad (e+d) = \frac{\text{width of grating}}{\text{total number of lines on the grating}} = \frac{3}{15000} cm,$

$$n = 1 \text{ and } \lambda = 5000\text{Å} = 5 \times 10^{-5} cm$$

$$\frac{3}{15000}\sin\theta = 1 \times 5 \times 10^{-5}$$

or $\quad \sin\theta = \frac{1 \times 5 \times 10^{-5} \times 15000}{3} = \frac{1}{4} = 0.25$

$$\theta = \sin^{-1}(0.25) = 14°29'$$

Ex.8. A plane transmission grating having 5000 lines per cm is being used under normal incidence of light. Answer the following questions:

a. What is the longest wavelength of light for which a spectrum can be seen?
b. What is the highest order spectrum that can be seen for the light of wavelength 6000Å?
c. The above-mentioned spectral line in the second order spectrum overlaps with another spectral line in next higher order. Find the wavelength of the other spectral line.
d. If the width of opaque parts be double than that of transparent parts of the grating then which orders of spectra will be absent?
e. If 90% of the width of the ruled width of the grating is covered, what will happen to the observed spectrum?

f. If the number of lines per cm is kept unchanged but the shape of grooves is changed, then how the dispersive power of the grating and the relative intensity of the spectra of different orders should be affected?

Sol. (a) For a plane transmission grating, we have

$$(e+d)\sin\theta = n\lambda$$

or $$\lambda = \frac{(e+d)\sin\theta}{n}$$

For a given grating i.e., for given value of $(e + d)$, λ will be maximum, when $\sin\theta$ is maximum and n is minimum.

Maximum value of $\sin\theta = 1$, and minimum value of $n = 1$.

$\therefore$ Maximum value of $\lambda = \lambda_{max} = \frac{(e+d)\times 1}{1} = (e+d)$

Here $$(e+d) = \frac{1}{5\times 10^3}\,cm = 2\times 10^{-4} cm.$$

$\therefore$ $$\lambda_{max} = 2\times 10^{-4} cm = 2\times 10^{4} Å$$

(b) From the relation, $(e + d)\sin\theta = n\lambda$, we have

$$n = \frac{(e+d)\sin\theta}{\lambda}$$

For maximum value of n, $\sin\theta = 1$

$$n_{max} = \frac{(e+d)}{\lambda} = \frac{2\times 10^{-4}}{6\times 10^{-5}} = 3.33 = 3(\text{whole number})$$

(c) If λ' is the desired wavelength, then

$$(e+d)\sin\theta = n\lambda = n'\lambda'$$

$\therefore$ $$\lambda' = \frac{n\lambda}{n'} = \frac{2\times 6000Å}{3} = 4000Å$$

(d) The direction of principal maxima is given by

$$(e+d)\sin\theta = n\lambda \quad (1)$$

Here e is the width of transparent space and d that of opaque space.

The direction of minima in the diffraction pattern due to a single slit is given by

$$(e+d)\sin\theta = m\lambda \quad (2)$$

Those principal maxima will be absent in the diffraction pattern which lie in the direction given by (2). Dividing (1) by (2), we get

$$\frac{(e+d)}{e} = \frac{n}{m}$$

or $$n = \frac{(e+d)}{e} m$$

Here $d = 2e$

$$n = 3,6,9, \ldots (\text{since } m = 1,2,3, \ldots .)$$

Thus 3rd, 6th, 9th, ... orders of spectra will be absent.

(e) If 90% of the width of the ruled width of the grating is covered, effective lines, i.e., number of slits will reduce to $\frac{9}{10}$ of its previous value i.e., now the number of slits will be 0.9N. However, the grating element remains the same. The following will be the changes observed in the spectrum:

(i) As the intensity is proportional to N^2, the intensity of spectral lines will become 0.81 times its previous value i.e., the lines will become fainter.

(ii) As the width of a spectral line is inversely proportional to N, the width of spectral lines become $\frac{10}{9}$ times its previous value i.e., the lines will become wider.

(f)If the grating element $(e + d)$ is kept unchanged, the dispersive power $\frac{d\theta}{d\lambda} = \frac{n}{(e+d)\cos\theta}$ for a given order would remain unchanged.

If the shape of the grooves is changed the relative intensity of different order spectra would change.

Ex.9. What is the highest order spectrum which may be seen with light of wavelength 5×10^{-5} cm. by means of grating with 3000 lines/cm.? Calculate the wavelengths in the visible spectrum $(3.5 \times 10^{-5}$ cm to 7.0×10^{-5} cm.) which coincide with the fifth order spectrum of this light.

Sol. (i) For plane transmission grating, we have

$$(e + d)\sin\theta = n\lambda$$

Or $n = \frac{(e + d)\sin\theta}{\lambda}$

Here n will be maximum, when $\sin\theta = 1$

$$n_{max} = \frac{(e + d)}{\lambda}$$

He$(e + d) = \frac{1}{3000}$ cm. and $\lambda = 5 \times 10^{-5}$ cm

$n_{max} = \frac{1/3000}{5 \times 10^{-5}} = \frac{1}{5 \times 10^{-5} \times 3000} = 6.6$ But the order can only be a whole number

$$n_{max} = 6.$$

(ii) If λ' is the wavelength in mth order, which coincides with wavelength λ in nth order, we have

$$(e + d)\sin\theta = n\lambda = m\lambda'$$

$$\lambda' = \frac{n}{m}\lambda$$

Here $n = 5, \lambda = 5 \times 10^{-5}$cm.

$$\lambda' = \frac{5 \times 5 \times 10^{-5}}{m} = \frac{25 \times 10^{-5}}{m}$$

If, $m = 3, \lambda' = \frac{25 \times 10^{-5}}{3} = 8.33 \times 10^{-5}$cm.

And if, $m = 4, \lambda' = \frac{25 \times 10^{-5}}{4} = 6.25 \times 10^{-5}$cm.

If, $m = 5, \lambda' = \frac{25 \times 10^{-5}}{5} = 5 \times 10^{-5}$cm.

If, $m = 6, \lambda' = \frac{25 \times 10^{-5}}{6} = 4.16 \times 10^{-5}$cm.

If, $m = 7, \lambda' = \frac{25 \times 10^{-5}}{7} = 3.57 \times 10^{-5}$cm.

If, $m = 8, \lambda' = \frac{25 \times 10^{-5}}{8} = 3.12 \times 10^{-5}$cm.

As λ' lies between 3.5×10^{-5} cm. and 7.0×10^{-5} cm, the wavelengths 6.25×10^{-5} cm., 4.16×10^{-5} cm, 3.57×10^{-5} cm. coincide with 5th order spectrum, of given light of wavelength 5×10^{-5} cm.

Ex.10 In a grating spectrum, which spectral line in fourth order will overlap with third order line of 5461Å ?

Sol. The wavelength λ' of m^{th} order will overlap with wavelength λ of n^{th} order, if

$(e + d)\sin\theta = n\lambda = m\lambda'$

or $\lambda' = \frac{n\lambda}{m}$

Here $n = 3, \lambda = 5461Å, m = 4$

$\lambda' = \frac{3\times5461}{4}Å$

$= \frac{16383}{4}Å \quad = 4095.75Å$

Ex.11 A diffraction grating used at normal incidence gives a green line $(5400Å)$ in a certain order superimposed on the violet line $(4050Å)$ of the next higher order. If the angle of diffraction is $30°$, how many lines per cm are there in the grating?

Sol. The directions of principal maxima for normal incidence are given by

$$(e + d)\sin\theta = n\lambda$$

If n^{th} order of wavelength λ_1 (say) is superimposed on $(n+1)^{th}$ order of λ_2, then we have

$(e + d)\sin\theta = n\lambda_1 = (n + 1)\lambda_2$

This gives $\frac{(e+d)\sin\theta}{\lambda_1} = n$ (1)

And, $\frac{(e+d)\sin\theta}{\lambda_2} = n + 1$ (2)

Subtracting (1) from (2), we get

$$(e + d)\sin\theta\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) = 1$$

or $$(e + d)\sin\theta\left(\frac{\lambda_1-\lambda_2}{\lambda_1\lambda_2}\right) = 1$$

or $$(e + d) = \frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}\cdot\frac{1}{\sin\theta}$$

Here $\lambda_1 = 5400Å = 5400 \times 10^{-8}cm., \lambda_2 = 4050 \times 10^{-8}cm., \theta = 30°$

$(e + d) = \frac{5400 \times 10^{-8} \times 4050 \times 10^{-8}}{5400 \times 10^{-8} - 4050 \times 10^{-8}} \times \frac{1}{\sin 30°}$

$= \frac{5400 \times 4050 \times 10^{-8}}{1350 \times \frac{1}{2}}$

$= \frac{5400 \times 4050 \times 2}{1350} \times 10^{-8}$

$\therefore$ Number of lines per cm $= \frac{1}{(e + d)}$

$$= \frac{1350}{5400 \times 4050 \times 2 \times 10^{-8}}$$

$$= \frac{1350 \times 10^8}{5400 \times 4050 \times 2} = 3086$$

Ex.12 The limits of visible spectrum are approximately 400 nm to 700 nm. Find the angular width of the first order visible spectrum produced by a plane diffraction grating having 15000 lines per inch when light is incident normally on the grating. (Given $\sin 13°40' = 0.237$ and $\sin 24°24' = 0.413$).

Sol. The directions of spectral lines are given by

$$(e + d)\sin\theta = n\lambda$$

or $$\sin\theta = \frac{n\lambda}{e+d}$$

Here $(e+d) = \frac{2.54}{15000}$ cm and $n = 1$ (for I order)

$$\sin\theta = \frac{1\times\lambda}{2.54/15000} = \frac{\lambda\times 15000}{2.54}$$

Let θ_1 and θ_2 be the angles of diffraction corresponding to
$\lambda = 400\text{ nm}(= 400\times10^{-9}\text{m} = 400\times10^{-7}\text{cm.})$
$\lambda = 700\text{ nm}(= 700\times10^{-9}\text{m} = 700\times10^{-7}\text{cm.})$
For first order, then

$$\sin\theta_1 = \frac{400\times10^{-7}\times15000}{2.54} = 0.236 \qquad \therefore \theta_1 = 13°40'$$

$$\sin\theta_2 = \frac{700\times10^{-7}\times15000}{2.54} = 0.413 \qquad \therefore \theta_2 = 24°24'$$

$\therefore$ Angular width of the first order visible spectrum $= \theta_2 - \theta_1$

$$= 24°24' - 13°40'$$
$$= 10°44'.$$

Ex.13 A plane transmission grating produces an angular separation of 0.01 radian between two wavelengths observed at an angle 30°. If the mean value of the wavelength is 5000Å and the spectrum is observed in the second order, calculate the difference in the two wavelengths.

Sol. The directions of spectral lines are given by

$$(e+d)\sin\theta = n\lambda \qquad (1)$$

Differentiating equation (1) we get

$$(e+d)\cos\theta\,\delta\theta = n\,\delta\lambda \qquad (2)$$

Dividing (2) by (1), we get

$$\frac{\cos\theta\,\delta\theta}{\sin\theta} = \frac{\delta\lambda}{\lambda}$$

Thus, the difference in two wavelengths $\delta\lambda$ is given by

$$\delta\lambda = \lambda\frac{\cos\theta}{\sin\theta}\,.\,\delta\theta = \lambda\cot\theta\,\delta\theta.$$

Here $\theta = 30°, \delta\theta = 0.01$ radian, $\lambda = 5000\text{Å} = 5000\times10^{-8}\text{cm}$

$$\delta\lambda = 5000\times10^{-8}\text{cm.}\times\cot30°\times0.01\text{ cm}$$
$$= 5000\times10^{-8}\text{cm}\times\sqrt{3}\times0.01\text{ cm}$$
$$= 5000\times10^{-8}\text{cm}\times1.732\times0.01\text{ cm}$$
$$= 86.6\times10^{-8}\text{cm} = 86.6\text{Å}$$

Ex.14 A plane transmission diffraction grating has 40,000 lines in all with grating element 12.5×10^{-5} cm. Calculate the maximum resolving power for which it can be used in the range of wavelength 5000Å.

Sol. The order of spectrum with a given grating is

$$n = \frac{(e+d)\sin\theta}{\lambda}$$

For maximum order $(\sin\theta)_{max} = 1$; therefore the maximum order that can be observed with a given grating of element $(e+d)$ for a wavelength λ is given by

$$n_{max} = \frac{(e+d)}{\lambda}$$

Here $(e+d) = 12.5\times10^{-5}$cm. and $\lambda = 5000\times10^{-8}$cm.

$$n_{max} = \frac{12.5\times10^{-5}}{5000\times10^{-8}} = 2.5$$

Thus, the second order is the highest which is observed with this grating.

Hence, maximum resolving power obtained = nN = 2 × 40,000 = 80,000

Ex.15 A diffraction grating is just able to resolve two lines of $\lambda = 5140.34 Å$ and 5140.85 Å in the first order. Will it resolve the lines 8037.20 Å and 8037.50 Å in the second order?

Sol. The resolving power of grating is given by

$$R.P. = \frac{\lambda}{d\lambda} = Nn$$

Here n is the order and N is the total number of lines on the grating.

According to first part of the problem

$$\lambda = \frac{(5140.34 + 5140.85)}{2} Å = 5140.595Å$$

$d\lambda = (5140.85 - 5140.34)Å = 0.51Å$

and $\quad n = 1$

Therefore, we have

$$R.P. = \frac{5140.595}{0.51} = 1 \times N$$

or $\quad N = \frac{5140.595}{0.51} = 10080$

The R.P. of the grating in second order = 2N = 2 × 10080 = 20160.

The resolving power required to resolve the lines 8037.50 Å and 8037.20 Å in the second order $= \frac{\lambda}{d\lambda}$.

Here $\quad \lambda = \frac{8037.20 + 8037.50}{5} = 8037.35Å$

and $\quad d\lambda = (8037.50 - 8037.20)Å = 0.30Å$

∴ R.P. required to resolve the two given lines

$= \frac{8037.35}{0.30} = 26791.17$

As the resolving power required (26791.17) is greater than the actual resolving power (20160) of the given grating; therefore, the given lines will not be resolved in the second order.

Ex. 16 A plane transmission grating has 16000 lines to an inch over a length of 5 inches. Find (a) the resolving power of the grating in the second order and (b) the smallest wavelength difference that can be resolved for light of wavelength 6000Å.

Sol. Number of lines per inch on the grating = 16000

length of ruled grating = 5 inches.

∴ Total number of lines on the grating = 16000 × 5 = 80000

wavelength, $\lambda = 6000Å = 6000 \times 10^{-8}$cm., n = 2

∴ resolving power = Nn

= 2 × 80000 = 1,60,000

The smallest resolvable wavelength difference $d\lambda$ is given by

$\frac{\lambda}{d\lambda} = Nn = 1,60,000$

$d\lambda = \frac{\lambda}{1,60,000} = \frac{6000 \times 10^{-8}}{1,60,000} = 375 \times 10^{-4}\text{cm} = 0.0375Å$

Ex.17 In the second order spectrum of a plane diffraction grating a certain spectral line appears at an angle of 10°, while another line of wavelength 5×10^{-9}cm. greater appears at an angle 3″ greater.

Find the wavelength of the lines and the minimum grating width required to resolve them. (Given $\sin 10° = 0.1736$ and $\cos 10° = 0.9848$)

Sol. The directions of maxima in a grating are given by

$$(e + d)\sin\theta = n\lambda \qquad (1)$$

Differentiating, we get $\quad (e + d)\cos\theta d\theta = n d\lambda \qquad (2)$

Dividing (1) by (2), we get

$$\frac{\sin\theta}{\cos\theta d\theta} = \frac{\lambda}{d\lambda}$$

Or $\lambda = \dfrac{\sin\theta\, d\lambda}{\cos\theta d\theta}$

Here $\quad \theta = 10°, d\theta = 3'' = \left(\frac{3}{60\times60}\right)^{\circ} = \frac{3}{60\times60} \times \frac{\pi}{180}$ radians

and $\quad d\lambda = 5 \times 10^{-9}\text{cm.} = 0.5 \times 10^{-8}\text{cm.}$

$$\therefore \lambda = \frac{\sin 10°}{\cos 10°} \times \frac{0.5\times10^{-8}}{\left(\frac{3}{60\times60}\times\frac{\pi}{180}\right)}$$

$$= \frac{0.1736 \times 0.5 \times 10^{-8} \times 60 \times 60 \times 180}{0.9848 \times 3 \times 3.14}$$

$$= 6063 \times 10^{-8}\text{cm.}$$

$$\lambda + d\lambda = 6063 \times 10^{-8} + 0.5 \times 10^{-8} = 6063.5 \times 10^{-8}\text{cm.}$$

Thus, the wavelengths of lines are 6063×10^{-8}cm. = 6063Å and 6063.5×10^{-8}cm. = 6063.5Å

The resolving power of the grating required to resolve the lines in the second order is given by

$$\frac{\lambda}{d\lambda} = Nn$$

Or $\dfrac{6063\times10^{-8}}{0.5\times10^{-8}} = N \times 2$

Or $N = \dfrac{6063\times10^{-8}}{0.5\times10^{-8}\times2} = 6063.$

Here, minimum grating width required = N. (e + d)

$$= N\frac{n\lambda}{\sin\theta} = 6063 \times \frac{2\times6063\times10^{-8}}{0.1736} = 4.2 \text{ cm.}$$

*****Review Questions and Problems

Based on Diffraction through a single slit

1. What do you mean by diffraction?
2. What is the difference between interference and diffraction?
3. Name the two classes of diffraction. What is the difference between them?
4. Distinguish between the Fresnel and the Fraunhofer classes of diffraction.
5. Explain, giving suitable examples, what is meant by the Fraunhofer class of diffraction phenomena.
6. Is there any fundamental difference between interference and diffraction? Give reasons.
7. Explain clearly the difference between Fresnel and Fraunhofer classes of diffraction phenomena. Discuss Fraunhofer diffraction at a single slit and show that the intensity of the first secondary maximum is roughly 4.5% of that of the principal maximum.
8. Write the conditions of maxima and minima in diffraction pattern due to single slit.

9. What is the ratio of relative intensities of successive principal maxima in diffraction due to a single slit?
10. What is the width of central maxima in the diffraction due to a single slit?
11. What happens to the diffraction pattern of a single slit if the slit is made narrower?
12. Discuss Fraunhofer type diffraction produced by a narrow single slit of width a and illuminated by monochromatic light of wavelength λ. Also deduce the positions of maxima and minima and plot the intensity distribution curve.
13. Discuss the phenomenon of Fraunhofer diffraction at a single slit and show that relative intensities of successive maxima are nearly $1 : \frac{4}{9\pi^2} : \frac{4}{25\pi^2} : \frac{4}{49\pi^2} : \ldots\ldots$

What is the nature of the pattern if the slit width is equal to λ ?

14.Discuss Fraunhofer diffraction due to a single slit. Extend the theory to the case of a plane transmission grating. Explain what is meant by diffraction spectra of different orders and state the condition under which the grating spectra of even order are absent.

15. Light of wavelength 5000Å is incident normally on a slit. The first minimum of the diffraction pattern is observed to lie at a distance of 5 mm. from the central maximum on a screen placed at a distance of 2 metres from the slit. Calculate the width of the slit. [Ans. 0.02 cm.]

16. A linear aperture whose width is 0.002 cm. is placed immediately in front at a lens of focal length 60 cm. If this aperture is illuminated by a beam of parallel rays whose angle of incidence is zero and whose wavelength is 5×10^{-5} cm. What will be the distance between the centre and the first dark band of the diffraction pattern on a screen placed 60 cm. from the lens?
[Ans. 1.5 cm.]

Based on Diffraction through double slit and circular aperture

17(a)Discuss the Fraunhofer diffraction at a double slit and deduce intensity distribution.
(b)Explain the effect of slit width separation on the diffraction pattern.
18. Describe the feature of a double slit Fraunhofer's diffraction pattern. What is missing order?
19.Discuss Fraunhofer diffraction due to double slit with necessary theory and discuss the intensity distribution. Why missing order occurs in this?
20.Describe the features of a double slit Fraunhofer's diffraction pattern. How does it change if the distance between the slit centres is varied, keeping the slit width constant?
21. Give the theory of Fraunhofer's diffraction at a circular aperture.
22.What are Airy's rings? Explain

Based on Diffraction grating

23. Give the construction and theory of plane transmission grating.
24. What do you mean by a diffraction grating?
25. What do you mean by the term principal maxima and secondary maxima in the case of diffraction pattern of a grating?
26. What is the effect of closeness of rulings and width of ruled surface on the grating spectrum?
27. What is the advantage of increasing the number of rulings in a grating?
28. Two plane diffraction grating A and B have same width of ruled surface but A has greater number of lines than in B. Which has greater intensity of fringes? Greater width of principal maximum?
29. (a) Give the construction and theory of plane transmission grating and explain the formation of spectra by it.

(b) Obtain an expression for the maximum number of orders available with a plane transmission grating.

30. Discuss the theory of a diffraction grating. Describe in detail how you would use a transmission grating to determine the wavelength of light.

31. Give the theory of a plane transmission grating and show how would you use it to find the wavelength of light?

32. Describe a plane transmission grating. Derive the formula connecting the wavelength of light, its deviation and grating constant. Hence obtain an expression for the wavelength of light in terms of angle of diffraction in the first order and the number of rulings on the grating.

33. If the opaque spaces are twice as wide as the transparent spaces, which orders of the spectra will be absent?

34. How many orders will be visible if the wavelength of incident radiation be 5000 Å and the number of lines on the grating be 2620 to an inch. [Ans. 19]

Based on Missing order or absent spectra

35. (a) What do you understand by missing order spectrum? What particular spectra would be absent if the width of transparencies and opacities of the grating are equal?

36. (b) If in a grating opaque space are exactly 2.0 times the transparent spaces, which order of spectra will be absent?

37. Explain the formation of spectra by a plane transmission grating. What do you understand by overlapping and absent spectra? Under what conditions only the first order spectrum will be seen?

38. (a) How do you account for the absence of some orders of spectrum obtained by plane transmission grating?

39. If the width of the transparent element of grating is equal to the thickness of line, which of the orders will disappear?

40. What is meant by diffraction of light? Which order of spectra would be absent if the width of the transparencies of the grating are equal

Based on Resolving and Dispersive power of grating

41. What do you mean by resolving power of an optical instrument?

42. Distinguish between the dispersive power and the resolving power of a diffraction grating. Deduce expressions for them.

43. Obtain an expression for the dispersive power of a grating. Why the grating spectrum is called a normal spectrum?

44. State Rayleigh's criterion of resolution.

45. Two gratings A and B, have the same width of the ruled surface but A has greater number of lines. Giving reasons compare in the two cases the (i) dispersive and (ii) resolving power, of gratings.

46. What do you understand by the term resolving power of a grating? Explain the Rayleigh criterion for the limit of resolution.

47. Two lines in the fourth order spectrum formed by a plane grating are just resolved. If the lines are due to wavelengths 5890 Å and 5896 Å, find the number of lines on the grating. [Ans. 246]

48. Find the minimum number of lines that a diffraction grating would need to have in order to resolve in first order the red doublet given by a mixture of hydrogen and deuterium. The wavelength difference is 1.8 Å at λ = 6563 Å. [Ans. 3646]

****

Unit – III

Polarization: Introduction to polarization, Brewster"s law, Malu's law, Nicol prism, double refraction, quarter- wave and half-wave plates, optical activity, specific rotation, Laurent half shade polarimeter.

Chapter 5

Polarisation of Light

Introduction

The electromagnetic wave character of light has been proved by Maxwell's electromagnetic theory. The theory explains that in the propagation of light, electric and magnetic fields oscillate perpendicular to each other and also perpendicular to the propagation of flow of energy. Unpolarised light has the oscillations of electric field vectors in all directions in the plane perpendicular to the propagation. If by any means all the electric field vectors are confined in one direction then the light is said to be polarized. Obviously, the magnetic field vectors will be oscillating perpendicular to the electric field Fig. 5.1.

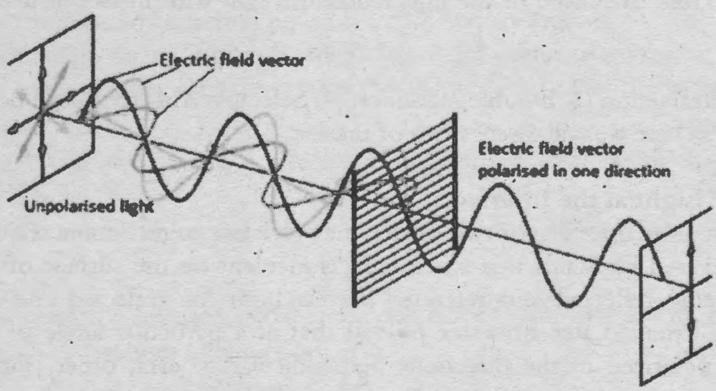


Fig. 5.1. Polarised electric field vector

There are a lot of mechanisms by which the electric vectors of un-polarised light are forced to oscillate in one direction. The oscillation of charge between two poles of an electric dipole is a basic principle for generation of polarized electromagnetic waves Fig. 5.2.

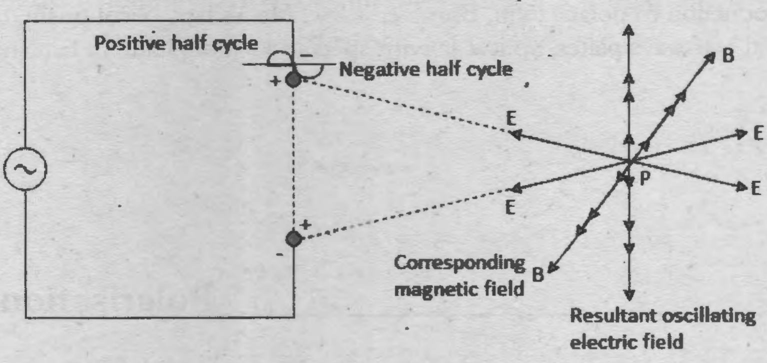


Fig.5.2. Production of oscillating electric and magnetic field using electric dipole

As light is a high frequency transverse electromagnetic wave, it is very difficult to make such type of di-pole antenna, which can produce polarized electromagnetic waves of visible range. But when the light interacts with the surface of a matter, the oscillations of atoms and molecules are produced in the matter in such a way that is required to produce transverse oscillating electric and magnetic field, giving rise to the production of polarized lights in reflection and transmission both. As these atoms or molecules oscillate with the forced frequency of the incident electromagnetic waves, the frequency does not change. Here, are some of the interactions of light with matter in which polarized light is produced.

(1) Reflection (2) Refraction (3) Double refraction (4) Selective Absorption (5) Scattering
As a part of syllabus here we will discuss two of these.

5.1 Interaction of Light at the Interface

Interaction of light at an interface of two media, may give rise to reflection, transmission or both. In 1808, it was discovered by Malus that when light is incident on the surface of any matter which is capable of producing reflected and refracted waves, both the reflected and refracted waves are polarized at some extent. Later Brewster proved that at a particular angle of incidence, both the waves are highly polarized in the directions perpendicular to each other (Fig.5.3). This angle of incidence is different for different interfaces, where the reflected and refracted waves are highly polarised (known as Brewster angle). It is also proved that in this situation sum of the angle of incidence and angle of refraction is 90° (Fig. 5.3).

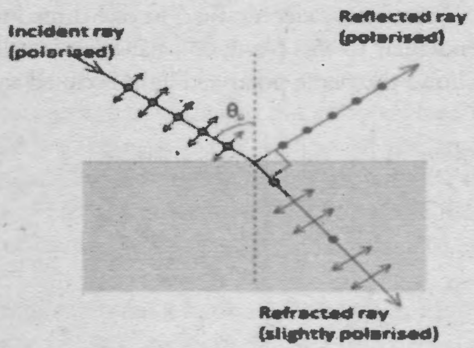


Fig. 5.3. Brewester's law at the polarising angle

Polarisation by reflection was well studied by Brewster and Malus. Brewster concluded about the angles of incidents of light of different interfaces for which the reflected and refracted waves are highly polarized. That is called Brewster's angle for a particular interface. Malus worked on the change in the intensity of polarised light when it encounters with a number of analysers (polarisers). Let us derive the formulations of these two laws.

5.2 Brewster's Law

Brewster performed a number of experiments to reach the conclusion that when unpolarised light is incident on a specific angle on an interface called Brewster's angle, the tangent of the Brewster's angle is always equal to the relative refractive index of the interface of two media. i.e.

$$\mu = \tan i_p \qquad (1)$$

where i_p is the angle of incidence for which an unpolarised light is highly polarized in both reflection and refraction.

But from Snell's law the relative refractive index of a transparent medium is given as:

$$\mu = \frac{\sin i}{\sin r} \qquad (2)$$

where i and r are the angles of incidence and refraction respectively. Brewster's angle of incidence i can be taken as i_p. So, the equation (2) becomes:

$$\mu = \frac{\sin i_p}{\sin r} \qquad (3)$$

from (1) and (3) we have

$$\mu = \tan i_p = \frac{\sin i_p}{\sin r}$$

or $\dfrac{\sin i_p}{\cos i_p} = \dfrac{\sin i_p}{\sin r}$

or $\cos i_p = \sin r$

$\Rightarrow (90° - i_p) = r$

or $i_p + r = 90°$

It means that whenever light is incident on Brewster's angle, the sum of the angle of incidence and the angle of refraction is 90°or reflected wave and refracted wave are perpendicular to each other.

5.3 Malus Law

As the polarized light has the oscillation of electric field in one direction, this plane is called as plane of vibration and a plane perpendicular to this plane of oscillation is called plane of polarization. The direction in a polarizer which allows the plane polarized light is called the axis of polarizer.

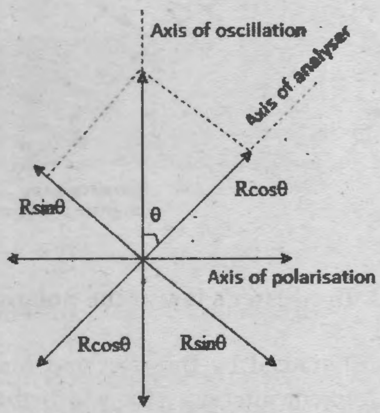


Fig. 5.4. Malus Law

Malus experimentally verified and later derived that if a plane polarized light falls on an analyser making an angle θ between the plane of vibration and the axis of analyser, the intensity passing through the analyser is proportional to the square of cosine of the angle i.e.

$I \propto \cos^2 \theta$

And it comes out as $I = I_0 \cos^2 \theta$, where I_0 is the intensity of polarized light incident on the analyser. For ideal polarizer, I_0 comes out to be nearly half of the intensity of un-polarised light. We can say;

$I = \frac{I_U}{2} \cos^2 \theta$, where I_U is called intensity of un-polarised light.

5.3.1 Derivation of Malus' Law

According to the Fig.5.4. as shown above, the amplitude of electric field of polarized light can be resolved along and perpendicular to the axis of analyser. The component of electric field along the axis of analyser is $R\cos \theta$. Thus, the intensity which passes through the analyser will be;

$$I \propto R^2 \cos^2 \theta$$

or $I \propto I_0 \cos^2 \theta$

Approximately we can write,

$$I = I_0 \cos^2 \theta$$

which is the Malus' Law.

5.4 Polarisation by Double Refraction

Introduction

The year 1669 marked as one of the greatest discoveries by a Dutch philosopher, Erasmus Bartholinous. He observed that whenever a ray of ordinary light falls on a crystal-like Iceland spar or calcite, it gets split into two refracted rays: one that obeys ordinary laws of refraction and the other does not obey the laws. Huygen's named this phenomenon as double refraction or a 'strange refraction' of calcite. He observed that both a slab of calcite and a slab of glass can lift the print underneath; glass lifting only one image, calcite on the other hand two images. Calcite thus exhibits an outstanding property of double refraction. In 1690, Huygen analyzed that both the refracted rays are to be linearly polarised in mutually perpendicular planes to each other. Later, it was also observed that the phenomena of double refraction was not seen in all the type of refractions but occurs in when light refracts through an anisotropic material. Now we will discuss in brief about isotropic and anisotropic materials, that will help to understand the Huygen's theory of double refraction i.e., Huygen's hypothetical theory of double wave front at the interface of an anisotropic material.

Isotropic Materials: The arrangement of atoms in isotropic materials is in such a manner in all directions such that most of the physical properties inside the material do not vary with directions. For example, electrical conductivity, thermal conductivity, velocity of sound, velocity of light, etc. are not direction dependent for isotropic materials. Obviously, the refractive index will be same along all the directions. Glass, water, air etc. are some of the examples as the isotropic materials from the point of view of optical properties.

Anisotropic Materials: If the arrangement of the atoms is in such a way that the various physical properties inside the material, are direction dependent then the materials are called anisotropic materials e.g., calcite, tourmaline, quartz, mica, topaz, aragonite etc

Refractive Property of an Anisotropic Material

When light interacts with an interface of an optically transparent medium, the electrons oscillate, due to electrical oscillations of the e.m. waves, under the constraints of forces of interaction between electrons and lattice. These oscillations are responsible for propagation of light inside the medium (refraction) and reflection. And the dipole like oscillation of an electron may be responsible for generation of polarized electromagnetic wave; which has to be discussed under the topic polarization by reflection and refraction. Obviously, the frequency of reflected and refracted light is same due to forced oscillations of electrons. If the medium is isotropic, the transfer of these electrical oscillations in all directions will be of exactly same type resulting in the direction independent optical properties inside the materials. But if the material is anisotropic, the transfer of electrical oscillations of an electron is not the same type in all the directions, resulting direction dependent optical properties of anisotropic materials.

5.4.1 Double Refraction

The phenomenon of double refraction is seen in crystals like calcite, tourmaline, quartz, mica, topaz, aragonite etc. These crystals are rigorously anisotropic in nature, as already mentioned above. It is now confirmed that the anisotropy of these crystals is responsible for the double refraction in these types of crystals. It has been verified that all the materials giving rise to double refraction, are anisotropic. The phenomena of double refraction discovered by Erasmus Bartholinous and later explained by Huygens, is a peculiar property of anisotropy of some type of crystals. As now we are going to discuss polarization by double refraction in detail, so let us first understand some important points about the geometry of doubly-refracting crystals, which will be helpful in understanding double refraction and ultimately the polarization by double refraction.

5.4.2 Uniaxial and Bi-axial Crystals

Whenever a ray of unpolarised light is made to fall on doubly refracting crystals, we get two refracted rays. Such crystals are of two types:

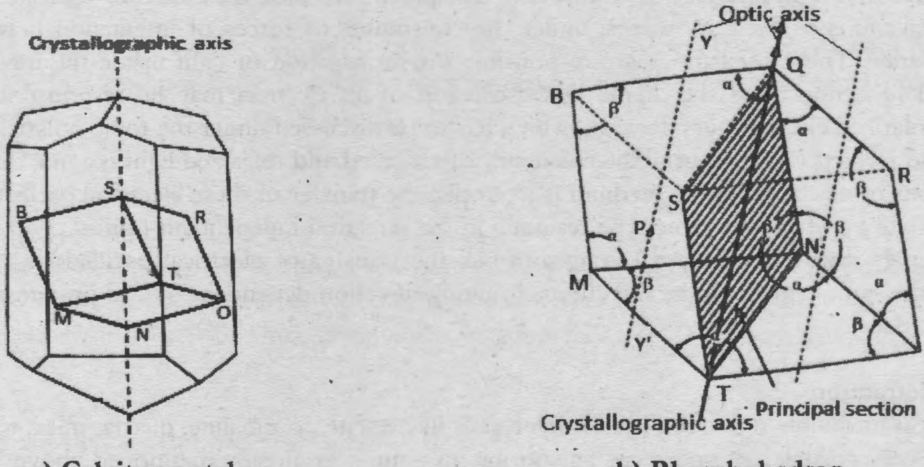
(i) Uniaxial crystals (ii) Biaxial crystals

Uniaxial crystals are the crystals having one direction, called optic axis, along which double refraction does not takes place or it may be said that two refracted rays travel with the same velocity along the optic axis. But if the unpolarised light falls in the direction other than optic axis of the uniaxial crystals, there are two refracted rays, one called ordinary and the other one is called extra-ordinary ray. The examples of some of the uniaxial crystals are calcite, tourmaline, quartz, etc.

Biaxial crystals are the crystals having two optic axis and none of the two refracted rays follow ordinary laws of refraction or both the rays are extraordinary rays. Topaz, aragonite etc. are some of the bi-axial crystals. Since, the study of uniaxial crystals is a bit simpler than the biaxial crystal and it is the topic of interest of current study, so we will focus on the most common example of this category i.e., calcite crystal.

5.4.3 Geometry of Calcite Crystal / Iceland spar

Iceland spar or hydrated calcium carbonate $CaCO_3$, times ago were abundantly found in Iceland as large transparent crystals.



a) Calcite crystal b) Rhombohedron

Fig. 5.5. Structure of calcite

Some of the properties of this crystal are:

(1) $CaCO_3$ can crystallize in a large colourless clear hexagonal prism with a blunt pyramid at each end as in Fig5.5(a.)

(2) It can be easily reduced by breakage or cleavage obliquely in three definite planes forming a rhombohedral body whose six rhombic sides makes an angle of $45^0 23'$ with the lines joining the vertices of the pyramids, the crystallographic axis of the crystal.

(3) The crystal is bounded by six faces each of which is a parallelogram with angles $\alpha = 101^0 55'$ and $\beta = 78^0 5'$.

(4) The three obtuse angles α meet at the two diametrically opposite corners Q and T as in Fig. 5.5(b), also known as blunt corners of the crystal. However, at the rest of the six corners there are two acute angles each equal to β and one obtuse angle α.

5.4.4 Ordinary and Extraordinary Rays

When an unpolarised light is incident normally on a calcite crystal it usually splits up into two refracted beams. At the exit face two rays recombine to make again ordinary unpolarised light.

The phenomenon of double refraction can be easily observed with an easy experiment: Mark an ink dot on a piece of paper and now place a calcite crystal on the dot, We can observe two images: one in which image remains stationary and the other image rotates with the rotation of crystal. The stationary image is called an ordinary image (O) and the second one, the extraordinary image (E) as shown in Fig. 5.6.

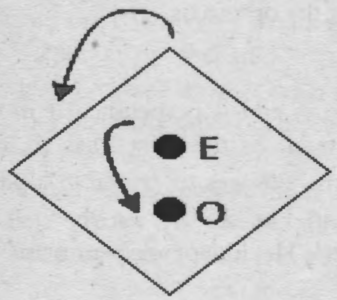


Fig.5.6. Two images E and O showing double refraction

Difference between the O ray and E ray

(i) The image O, lies in the direction of incident beam as it follows the ordinary laws of refraction. In contrast to it, the other image E is found to be separated from the O image despite perpendicular incidence as it follows extraordinary laws of refraction behaving in an extraordinary manner.

(ii) The velocity of O ray is the same in all directions within the crystal as the ratio sin i /sin r (which physically represents the ratio of the velocity of light within vacuum to that in the refracting medium) is constant within the crystal and so is the refractive index. For an extraordinary ray, the ratio sin i/sin r varies with the angle of incidence, its velocity is different in different directions within the crystal and hence the refractive index is not constant along the different directions.

(iii) O ray is always polarized in the plane of incidence but E ray is polarized in the perpendicular direction to the plane of incidence.

5.4.5 Axis of Crystal / Crystallographic Axis

The two diagonally opposite solid angles of calcite are formed by intersection of three obtuse angles of three faces. A line into the crystal at one of these corners and equally inclined to three faces is called a crystallographic axis. For a crystal having equal edges, as shown in Fig.5.5 (b), the shortest diagonal QT is the crystallographic axis.

5.4.6 Optic Axis

A line passing through any one of the blunt corners and making equal angles with each of the three edges which meet here gives the direction. This particular direction is thus named as "the optic axis", and is determined by the crystallographic axis. In fact, any direction in the crystal parallel to the crystallographic axis is called an optic axis. There is only one direction in the calcite crystal through which when a ray of light is incident, the O ray and E ray does not separate. Both of these rays travel with the same velocity along this direction. This optic axis is thus not a line but a direction. Through any given point P within the crystal, only one line YY′, can be drawn parallel to the crystallographic axis. This line gives the direction of the optic axis for this point and for all the points lying on it. If a ray of light is incident along the optic axis or in a direction parallel to it, then there is no splitting up into two rays. It is only when the light is incident other than the optic axis, the phenomenon of double refraction is observed.

Note: Optic axis is not obtained merely by joining the blunt corners. In a special case only when the three edges of the crystal are equal, the line joining the two blunt corners coincides with the axis of the crystal and gives the direction of the optic axis.

5.4.7 Principal Section

A plane which contains the optic axis and is perpendicular to two opposite faces of the crystal is called a principal section of the crystal. As the crystal has six faces, thus for every point there are three principal sections. This cuts the surfaces of crystal in a parallelogram with angles of 109° and 71°, as illustrated in for example in Fig. 5.5 (b) above, QSTN, represents the principal section through the blunt edges of the crystal. The independent principle section is also shown in Fig. 5.7.

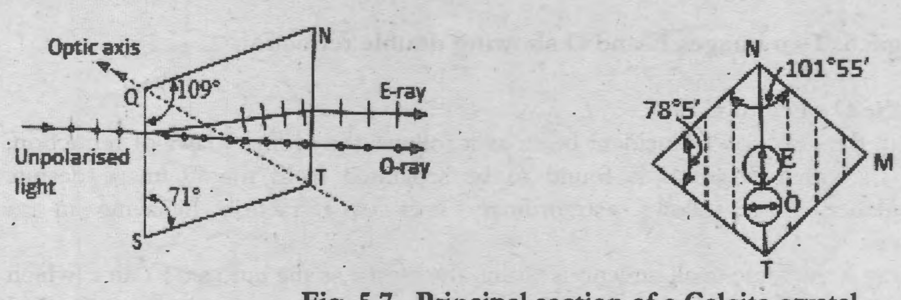


Fig. 5.7. Principal section of a Calcite crystal

5.4.8 Double Refraction Theory in Uniaxial Crystals – Huygen's, Explanation

Huygen's, after the discovery of the phenomenon of double refraction in uniaxial crystals like calcite, extended his explanation on the same, based on the principle of secondary wavelets. He was keen to explain the strange properties of calcite by his theory with the help of some postulates. Huygen's, postulates of double refraction are as follows:

1. An incident wavefront disturbs every point of the crystal giving rise to the origin of two wavelets, one the ordinary and the other extraordinary wavelet so as to understand the ordinary and extraordinary refracted ray within the crystal.
2. Both the laws of refraction are obeyed by an ordinary ray. Therefore, the refractive index of calcite for the ordinary ray can be given as:

$$\mu_o = \frac{\sin i}{\sin r_o} = \frac{v}{v_o}, \quad \text{always constant}$$

where, i = angle of incidence when an incident beam falls on the calcite crystal.

r_o = angle of refraction for an ordinary ray.

v_o = velocity of an ordinary ray.

Thus, refractive index of an ordinary ray comes out to be constant as it obeys the ordinary laws of refraction. It is thus independent of the direction of its propagation within the crystal as the velocity v_o of the ordinary wave is same in all the directions giving rise to a spherical wavefront.

3. The E ray does not obey ordinary laws of refraction. The refractive index of calcite for the extra-ordinary ray can be given as:

$$\mu_E = \frac{\sin i}{\sin r_E} = \frac{v}{v_E}, \quad \text{always changes with the direction}$$

r_E = angle of refraction for an extraordinary ray.

v_E = velocity of an extraordinary ray.

The refractive index of extraordinary ray is not constant in all directions, as it does not obey the ordinary laws of refraction and depends on the angle of incidence. It is dependent on the direction of its propagation within the crystal as the velocity v_E of the extra-ordinary wave is different in different directions giving rise to an ellipsoid wavefront.

4. In case of uniaxial crystals, the properties are perfectly symmetrical about it. Hence, the propagation of E ray takes place with an equal velocity in the directions which are equally inclined to optic axis. The ellipsoid wavefront is symmetrical about the optic axis. The Huygen's hypothesis concludes that the axis of symmetry or the axis of revolution of the ellipsoid coincides with the direction of optic axis through the point of origin of the wavelets.

5. No double refraction is observed along the optic axis in case of uniaxial crystals thereby leading to the equal velocities of the O and E rays along it. At the ends of the optic axis, the spherical wavelet and the ellipsoidal wavelet touch each other. In other words, the two wavelets touch each other at the points where the optic axis through the point of origin of wavelets intersects them.

Note: From Huygen's theory we can summarise that every point of the doubly refracting uniaxial crystal when disturbed by the incident wave can be considered as the source of two secondary wavelets: sphere for the ordinary wave and an ellipsoid for the extraordinary wave. The direction of the line joining the two points where the two wavefronts touch each other is called as optic axis.

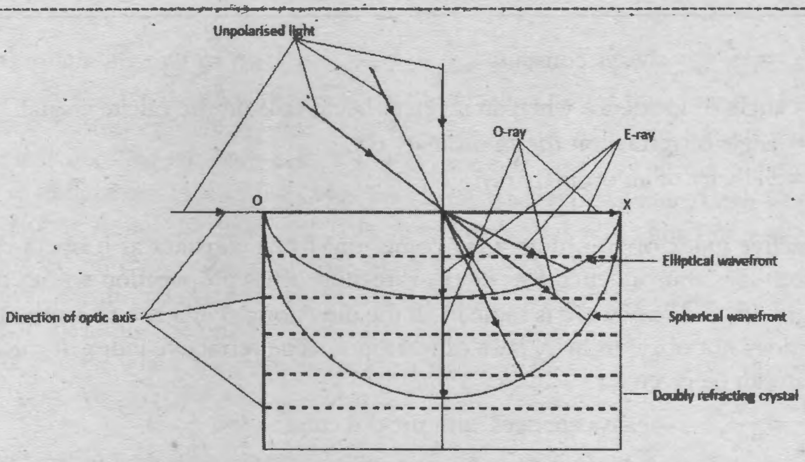


Fig. 5.8. Huygen's spherical and elliptical wavefronts for O- ray and E-ray for positive crystals

On the basis of experimental observations, he hypothesized the production of spherical and elliptical wavefronts for quartz and calcite crystals as shown in the Fig.5.8 and Fig.5.9 respectively. It can be seen from the Fig. 5.8 that the velocity of O-ray and E-ray are same along the optic axis and the velocity of E-ray decreases with respect to O-ray as the angle between incident light and optic axis increases. It is maximum along the optic axis and minimum when light is incident perpendicular to the optic axis. The ray originated from spherical wavefront i.e., O-ray does not have any variation in velocity with change in angle of incidence with the optic axis. These observations were obtained with quartz crystal. On the other hand, in case of a calcite crystal Fig.5.9, the velocity of E-ray increases as the angle between the incident ray and optic axis increases and it is maximum when light is incident perpendicular to the optic axis.

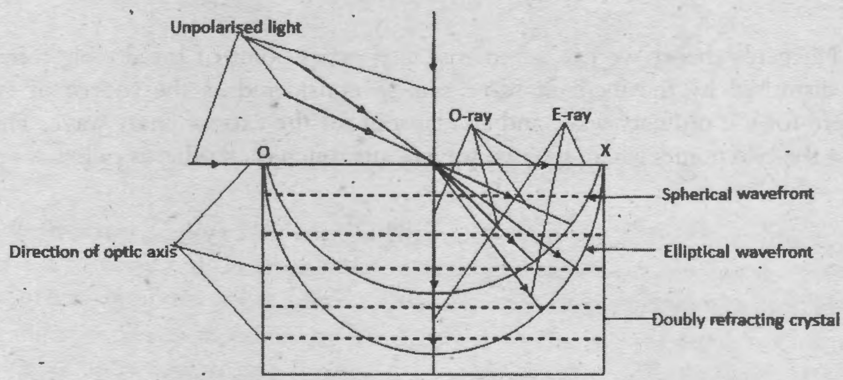


Fig. 5.9. Huygen's spherical and elliptical wavefronts for O- ray and E-ray for negative crystals

Here it should be noted that just for the name sake, the crystals in which the velocity of E-ray is minimum (equal to velocity of O-ray) along optics axis and maximum perpendicular to the optic axis are called as negative crystals, e.g. calcite crystal and in which velocity of E-ray is maximum (equal to velocity of O-ray), along optic axis and minimum in the direction perpendicular to the optic axis are called as positive crystals, i.e. Quartz.

On the basis of the observations obtained and according to the Huygens's hypothetical wavefront theory, we can formulate some optical properties like refractive index and velocities of E-ray and O-ray of negative and positive crystals.

As refractive index of light in a medium w.r.t vacuum/air =

$$\frac{\text{velocity of light in vacuum}}{\text{velocity of light in medium}}$$

So, the refractive index for O-ray will be given as;

$$\mu_0 = \frac{\text{velocity of light in vacuum}}{\text{velocity of the O ray in medium}}$$

Similarly, the refractive index for E ray will be;

$$\mu_E = \frac{\text{velocity of light in vacuum}}{\text{velocity of the E ray in medium}}$$

where μ_0, μ_E are the refractive indices for O-ray and E-ray respectively.

As the velocities of E-ray and O-ray is same along optic axis and different along other directions relative to the optic axis. Thus, on the basis of above discussion we can say:

$v_0 = v_E$ along optic axis for both negative and positive crystals.

$v_E > v_0$ in negative crystals for all directions except along optic axis.

$v_E < v_0$, in positive crystals for all directions except along optic axis.

Here v_0, v_E stands for velocity of O-ray and E-ray respectively.

The refractive indices for O-ray and E-ray can be related to each other as below:

$\mu_0 = \mu_E$ along optic axis for both negative and positive crystals

$\mu_E < \mu_0$, in negative crystals for all directions except along optic axis

$\mu_E > \mu_0$, in positive crystals for all directions except along optic axis

It can be concluded that all the optical properties of O-ray do not vary with the direction that is why it is called ordinary ray of light and on the other hand E-ray has the direction dependent optical properties, so it is termed as extra ordinary ray of light.

5.5 The Nicol Prism

Nicol prism was invented by William Nicol in 1828. It is a device made up of a calcite crystal and widely being used for producing and analysing plane polarised light.

5.5.1 Principle

When an ordinary unpolarised light passes through a crystal of calcite, two completely plane polarised beams, the O wave and the E wave, are obtained with their vibrations in two mutually perpendicular planes. O-ray or the ordinary ray have vibrations perpendicular to the principal section of the crystal and the E-ray or the extra–ordinary ray having vibrations parallel to the principal section of the crystal. William Nicol separated the ordinary beam from extra ordinary ray, by the phenomenon of total reflection, using special geometry of the doubly refracting crystal called Nicol prism to get a plane polarized light.

5.5.2 Construction

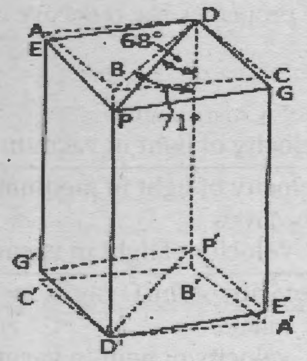


Fig. 5.10. Geometry of a Nicol prism

The Nicol prism is not an ordinary optical prism but instead a parallelopipedon. It is a calcite rhombohedron produced by cleavage from a clear natural crystal with length three times its width. From Fig.5.10, we can clearly see that the three obtuse angles of the three faces meet at the corners B and B′ of the crystal. Here, the principal section of the crystal (BDB′D) is formed by an imaginary plane passing through the edges BD′ and B′D which contains the optic axis. The natural crystal is grounded such that the angles BDB′ and B′D′B are equal to 71° in the crystal. The end faces ABCD and A′B′C′D′ are cut artificially by planes perpendicular to the principal section BDB′D′ so as to form new faces PGDE and P′G′D′E′ so that the angles PDP′ and P′D′P are equal to 68°. The crystal is then cut along a plane passing through P and P′ and perpendicular to the shorter diagonal of the end faces PD and P′D′ such that it becomes perpendicular to the principal section of the crystal. With the help of a thin layer of transparent adhesive, Canada Balsam (CB), these cut surfaces are cemented together. The refractive index of Canada Balsam lies in between the indices of refraction of calcite for the O and the E rays. For sodium yellow light at mean wave-length λ = 5893 **Å**, the values of refractive indices are as;

Index of calcite for O ray ($\mu_o = 1.65836$),

Index of Canada balsam ($\mu_{CB} = 1.55$)

Index of calcite for E ray ($\mu_E = 1.48641$)

$\Rightarrow \mu_o > \mu_{CB} > \mu_E$

This implies that Canada balsam is optically denser than calcite for E ray; as $\mu_{CB} > \mu_E$, and optically less dense than calcite for O ray as $\mu_{CB} < \mu_o$.

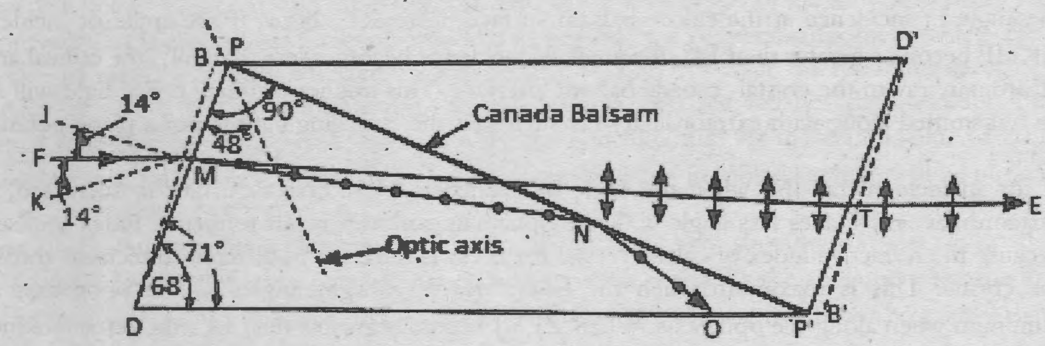


Fig.5.11. Action of Nicol prism

5.5.3 Working

Let us study that how Nicol prism produces the polarized light by eliminating one of the two refracted rays. Consider an unpolarised monochromatic ray FM entering the prism through the face PGDE parallel to the long sides of the prism as seen in Fig. 5.11. After entering through the prism, it gets splits up into an ordinary and extraordinary ray such that the extraordinary ray makes a larger angle with the optic axis (paths MN, MT). Both of them are plane polarized in such a way that the vibrations of O ray being perpendicular to the principal section of the crystal while that of E ray being in the principal section.

The ordinary ray traverses from prism i.e an optically denser medium ($\mu_o = 1.66$) to Canada balsam i.e. an optically rarer medium ($\mu_{CB} = 1.55$).The refractive index of ordinary ray with respect for crystal - Canada balsam interface is $\frac{1.66}{1.55}$. This ray is totally internally reflected for angles of incidence on the balsam greater than the critical angle, viz.

$$\phi = \sin^{-1}\frac{\mu_{CB}}{\mu_o} = \frac{1.55}{1.65} = 0.933$$

$$\phi = \sin^{-1}(0.933) = 69°$$

After total reflection, this ray traces the path NO and is finally absorbed by the blackened surface DP′. On the other hand, however the E-ray emerges out from the prism parallel to the incident ray as it is passing from an optically rarer medium (calcite) into an optically denser medium (Canada balsam) and hence is transmitted through the prism. Thus, the E-ray is plane polarized with vibrations parallel to the principal plane of the crystal and a highly plane polarized light is obtained.

By understanding the above explained mechanism, the Nicol prism can be used as a polarizer by eliminating one of the two refracted rays, using the phenomena of total internal reflection. But this phenomenon puts some limitations for using the Nicol prism for the polarization purpose.

5.5.4 Limitations

In the visible spectrum, Nicol prism acts as a very good polariser. But with too convergent or too divergent incident beams, they are ineffective in their working due to the limitations imposed by the phenomenon of total internal reflection. This can be explained as follows;

(i) If the angle of incidence for an incident ray KM, at the crystal surface is increased as in Fig.5.11., the angle of incidence at the calcite-balsam surface decreases. Now, if the angle of incidence ∠KMF becomes greater than 14°, the angle of incidence becomes less than 69°, the critical angle of ordinary ray to the crystal- canada balsam interface. This implies ordinary ray of light will also be transmitted along with extraordinary ray such that the emerging light is not a plane polarized light.

(ii) If for an incident ray JM, when the angle of incidence at the crystal surface is decreased, the extraordinary ray makes less angle with the optic axis and as a result refractive index increases, because the refractive index of calcite crystal for E ray is different in different directions through the crystal. This is maximum when the E ray travels at right angles to the optic axis and minimum when along the optic axis. When ∠FMJ becomes greater than 14°, the refractive index of the calcite crystal for E ray becomes greater than that of refractive index of canada balsam and at the same time the angle of incidence for E ray at the calcite-balsam becomes greater than the corresponding critical angle. Thus, E ray is totally reflected along with O ray and no light emerges out then.

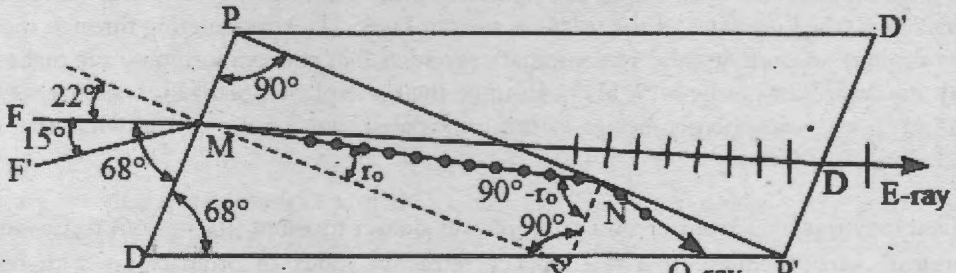


Fig. 5.12.

Let an incident ray travel along FM (see Fig.5.12) with an angle of incidence 22°; the corresponding angle of refraction for O ray may be given as:

$$\sin r_o = \frac{\sin i}{\mu_o} = \frac{\sin 22°}{1.65} = \frac{0.3746}{1.65} = 0.2270$$

$$r_o = \sin^{-1}(0.2270) = 13° \text{ approximately}$$

Hence, the angle of incidence for calcite-Canada balsam interface for O-ray will be;

$$\angle MNX = 90 - r_o = 77°$$

This is greater than the critical angle (69°) for the O ray. Now, if the angle of incidence be increased by 15°, then the corresponding angle of refraction r'_o is given by

$$\sin r'_o = \frac{\sin i'}{\mu_o} = \frac{\sin(22+15)}{1.65} = \frac{0.6018}{1.65} = 0.3647$$

$$r'_o = \sin^{-1}(0.3647) = 21.4°$$

The corresponding angle, $\angle MNX = 90 - r'_o = 68.6°$, is less than the critical angle 69°.

Therefore, the O ray is transmitted through balsam along with E ray. In the same manner, it can also be proved that if the light is incident beyond 14° towards the perpendicular on the face, the E-ray will also be totally internally reflected. This is because beyond this angle the refractive index of E-ray changes in such a way that the crystal - canada balsam interface satisfies the condition for total internal reflection for E-ray. Hence, so forth we can conclude that the maximum value of $\angle FMF'$ is about 14°, beyond which the O ray will also be transmitted and no polarized light will be obtained.

Therefore, in order to get a plane polarized light, the incident light should always be confined within 14° on its either side of a ray parallel to longer side of the crystal or the angle between the extreme rays of light incident on Nicol should be limited to about 28°. Hence, highly convergent or divergent light should not be used.

5.6 Production and Analysis of Polarised Light

Nicol Prism is widely being used for producing the plane polarized light. But when two plane polarized lights interfere, the resultant light may be plane polarized, elliptically polarized or circularly polarized light. In the next section, the analytical treatment of interference of two plane polarized light will help us to understand the production of plane polarized, elliptically polarized or circularly polarized light.

5.6.1 Interference of Two Plane Polarised Waves-Polarised in Mutually Perpendicular Planes

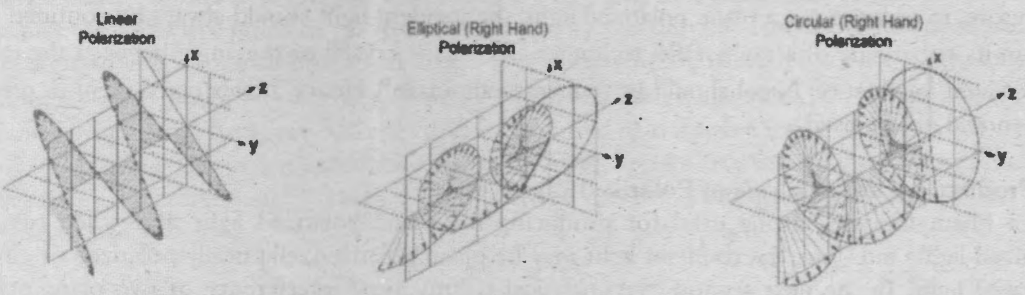
We have studied the interference of two light waves in Young's double slit experiment. Here, the basis of interference phenomena is same but the only difference is that in this case two light waves are polarized and plane of polarization of each wave is perpendicular to each other. In case of Young's interference, the resultant intensity does not vary if analysed with an analyser. But in case of interference of polarized light, the intensity may or may not vary on analyzing with an analyser.

Further, it is noted that in some cases, the resultant intensity variation analysing through an analyser, is such that it exactly resembles with a plane polarized light. In some other cases, the intensity of resultant light on passing through an analyser traces an ellipse with the angle of rotation of axis of the analyser and in some it looks as if it is unpolarised light. These three cases are termed as plane polarized, elliptically polarized and circularly polarized light respectively.
In the next section, we will discuss mathematical analysis of superposition of two plane polarized light, both polarized in perpendicular plane to each other. And further will derive the cases for plane polarized (linearly polarised), elliptically polarized and circularly polarized light. Let us define plane polarized (linearly polarised), elliptically polarized and circularly polarized light and discuss how a circularly polarized light differs with an unpolarised light.

Linearly Polarised Light: If the tip of the resultant of two plane polarized light traces a line with time then the resultant light is said to be linearly polarized. This linearly polarized light lies in a plane parallel to the propagation of light, that's why it may also termed as plane polarized light Fig.5.13(a).

Elliptically Polarised Light: If the tip of the resultant of two plane polarized light traces an ellipse when analysed with a rotating analyser is called an elliptically polarized light. Although, it is the locus of tip of the resultant, which is rotating in elliptical path, yet in most of the case it looks as permanently elliptical light on the screen, due to persistency of human eye as seen in Fig. 5.13 (b).

Circularly Polarised Light: If the tip of resultant of two plane polarized light traces a circular path when analysed with a rotating analyser is called circularly polarized light. On analyzing with the analyser it looks as unpolarised light yet it differs with it as unpolarised light have oscillations of electric field in all the directions perpendicular to the propagation of light and circularly polarized, is due to the locus of the path traced by the tip of the resultant of two plane polarized light as in Fig. 5.13(c).



a)Linear **b) Elliptical** **c)Circular Polarization**

Fig5.13.

5.6.2 Analysis of Superposition of Two Plane Polarized Light

Let there are two plane polarized progressive light waves, propagating along x-axis. One polarized along y-axis and other in z-axis. The equations of these can be written as;

$$E_y = E_{y_0} \cos(kx + \omega t) \tag{4}$$

$$E_z = E_{z_0} \cos(kx + \omega t + \delta) \tag{5}$$

Where each symbols used have their usual meaning.

On superposition, the resultant is given as the vector addition of two simple harmonic motions perpendicular to each other. i.e:

$$E = E_y + E_z$$
$$= E_{y_0} \cos(kx + \omega t) + E_{z_0} \cos(kx + \omega t + \delta)$$

As the propagation of resultant light wave will be along the same direction as the direction of each wave, we can ignore kx to make mathematical calculations simple. The equations (4) and (5) can be written as;

$$E_y = E_{y_0} \cos \omega t \tag{6}$$

$$E_z = E_{z_0} \cos(\omega t \pm \delta) \tag{7}$$

or

$$E_z = E_{z_0} (\cos \omega t \cos\delta + \sin \omega t \sin\delta)$$

$$= E_{z_0} (\cos \omega t \cos\delta + \sqrt{1 - \cos^2 \omega t} \sin\delta) \tag{8}$$

And the resultant equation of motion of the wave will be

$$E = E_{y_0} \cos\omega t + E_{z_0} (\cos\omega t \cos\delta + \sqrt{1-\cos^2\omega t}\,\sin\delta)$$

The shape of the curve traced by the tip of the resultant of two wave vectors can also be found by developing the relation between E_y and E_z with the change in phase difference between the two waves. Using (4), let us put $\cos\omega t = \frac{E_y}{E_{y0}}$, in equation (8) we get:

$$E_z = E_{z_0} \left(\frac{E_y}{E_{y0}}\right)\cos\delta + E_{z_0} \sqrt{1-\left(\frac{E_y}{E_{y0}}\right)^2}\sin\delta)$$

or

$$\{E_z - \left(\frac{E_{z0}}{E_{y0}}\right) E_y \cos\delta\} = E_{z_0} \sqrt{1-\left(\frac{E_y}{E_{y0}}\right)^2}\sin\delta)$$

On squaring both the sides,

$$\{E_z - \left(\frac{E_{z0}}{E_{y0}}\right) E_y \cos\delta\}^2 = \{E_{z_0} \sqrt{1-\left(\frac{E_y}{E_{y0}}\right)^2}\sin\delta\}^2$$

On simplification and rearranging the terms we get,

$$\frac{E_y^{\,2}}{E_{y0}^{\,2}} + \frac{E_z^{\,2}}{E_{z0}^{\,2}} - 2\frac{E_y E_z}{E_{y0} E_{z0}} \cos\delta = \sin^2\delta \qquad (9)$$

The equation (9) represents a general ellipse in YZ plane. Mathematically, we can say that the resultant of two simple harmonically oscillating vectors is an ellipse and the line and circle are the special cases of ellipse. In case, if the minor axis of the ellipse is zero, it will look as a line and if both the axis of the ellipse is equal, it is a circle.

Hence, if two plane polarized lights interfere with a general value of phase difference between them, the tip of the resultant amplitude of two waves will rotate tracing an ellipse giving rise to the production of elliptically polarized light. The exact shape and inclination to the coordinate axis can be found by putting the exact phase difference between the two interfering polarized lights.

5.6.3 Plane, Circularly and Elliptically Polarized Light

Here are some special cases, which can explain the production of plane, circularly and elliptically polarized light in details.

(A) Production of Plane Polarized light

(a) If $\delta = 2n\pi$ where $n = 0,1,2,3\ldots\ldots$ etc, equation (9) reduces to

$$\frac{E_y^{\,2}}{E_{y0}^{\,2}} + \frac{E_z^{\,2}}{E_{z0}^{\,2}} - 2\frac{E_y E_z}{E_{y0} E_{z0}} = 0$$

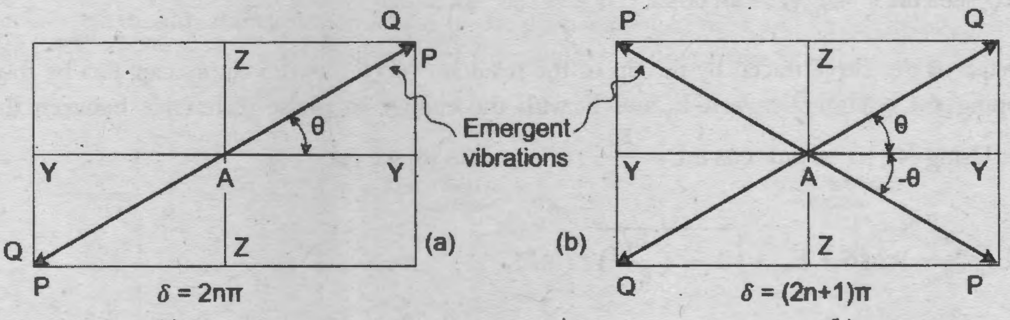
or

$$\left\{\frac{E_y}{E_{y0}} - \frac{E_z}{E_{z0}}\right\}^2 = 0$$

or

$$E_y = \pm \frac{E_{y0}}{E_{z0}} E_z \qquad (10)$$

Thus, resultant relation between E_y and E_z is a pair of coincidence straight lines with slope as, $m = \frac{E_{y0}}{E_{z0}}$ with y-axis. So, the resultant light will be plane polarized light as shown in Fig.5.14 (a)



(a) (b)

Fig.5.14 Resultant plane polarised lights corresponding to (a) $\delta = 2n\pi$ and (b) $\delta = (2n + 1)\pi$

(b) If $\delta = (2n - 1)\pi$ where n = 1,2,3….. etc
equation (9) reduces to

$$\frac{E_y^{\,2}}{E_{y0}^{\,2}} + \frac{E_z^{\,2}}{E_{z0}^{\,2}} + 2\frac{E_y E_z}{E_{y0}} = 0$$

or

$$\left\{\frac{E_y}{E_{y0}} + \frac{E_z}{E_{z0}}\right\}^2 = 0$$

or

$$E_y = \mp \frac{E_{y0}}{E_{z0}} E_z \tag{11}$$

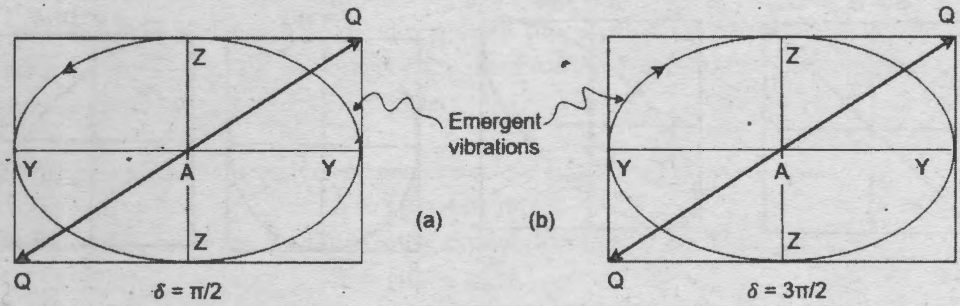
Thus, resultant relation between E_y and E_z is again a pair coincidence straight lines with a slope of $m = -\frac{E_{y0}}{E_{z0}}$ with y-axis. So, the resultant light will be plane polarized light as shown in Fig.5.14 (b).

(B) Production Of Elliptically Polarized Light

(a) If δ is an odd multiple of $\pi/2$ i.e., $\delta = (2n + 1)\pi/2$, the product term in equation (9) reduces to

$$\frac{E_y^{\,2}}{E_{y0}^{\,2}} + \frac{E_z^{\,2}}{E_{z0}^{\,2}} = 1$$

This is the equation of an ellipse, where the minor and major axis i.e., E_{y_0} and E_{z_0} coincides with the coordinate axis. So, if the interfering light waves are polarized and the phase difference between them is odd multiple of $\pi/2$, the resultant light will emerge as elliptically polarized light. Fig.5.15(a) for n = 0 and for n = 1, the rotation of electric vector will be like Fig.5.15 (b).



(a) (b)

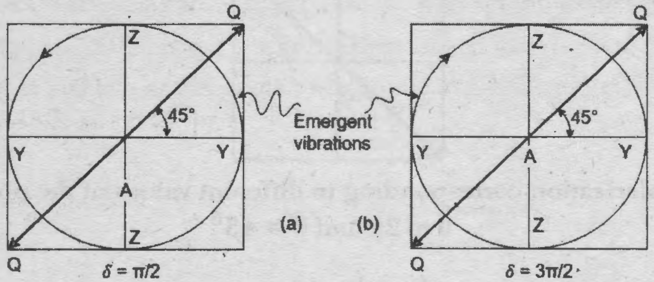
Fig.5.15. Resultant elliptically polarised light corresponding to (a) $\delta = \pi/2$ and (b) $\delta = 3\pi/2$

(C) Production of Circularly Polarized Light

If δ is an odd multiple of $\pi/2$ i.e., $\delta = \frac{(2n+1)\pi}{2}$ and minor and major axis of the ellipse i.e., E_{y_0} and E_{z_0} are equal, then equation (9) reduces to

$$E_y{}^2 + E_z{}^2 = E_{y_0}{}^2 \qquad (\text{let } E_{y_0} = E_{z_0} = E_{y_0} \text{ let}) \text{ and } \theta = 45^0$$

This is an equation of a circle with radius E_{y_0} . The orientation of rotating vector will be as in Fig.5.16(a) and 5.16(b) for phases $\delta = \pi/2$ and $\delta = 3\pi/2$ respectively.



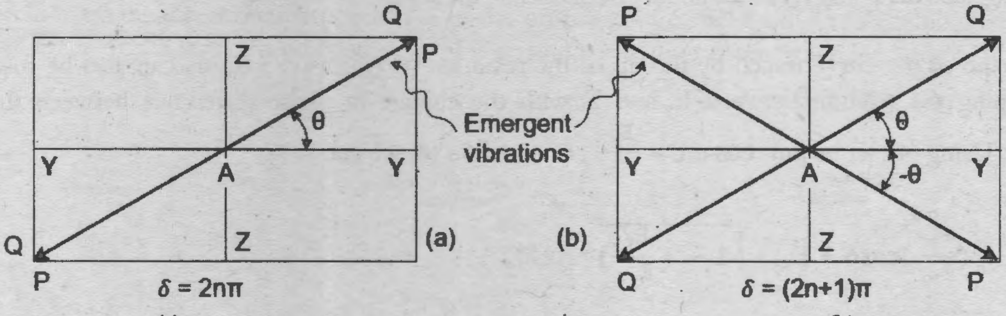
(a) (b)

Fig.5.16. Resultant circularly polarised light corresponding to (a) $\theta = 45^0$ and (b) $\delta = (2n + 1)\pi/2$

Here, we have discussed three cases in particular of interference of plane polarized light, depicting three particular shapes of emergent light. However, we can find the direction of rotation of resultant vector (clockwise or anticlockwise), slopes of plane polarized light with the axis and orientation or inclination of ellipse with axis etc. with any phase i.e., from 0 to 2π between two interfering plane polarized lights. Here are a few of them shown in Fig.5.17.

(i) Phase as 0° ii) Phase as $\frac{\pi}{4}$ iii) Phase as $\frac{\pi}{2}$ iv) Phase as $\frac{3\pi}{4}$

Thus, resultant relation between E_y and E_z is a pair of coincidence straight lines with slope as, $m = \frac{E_{y_0}}{E_{z_0}}$ with y-axis. So, the resultant light will be plane polarized light as shown in Fig.5.14 (a)



(a) (b)

Fig.5.14 Resultant plane polarised lights corresponding to
(a) $\delta = 2n\pi$ and (b) $\delta = (2n + 1)\pi$

(b) If $\delta = (2n - 1)\pi$ where n = 1,2,3….. etc
equation (9) reduces to

$$\frac{E_y{}^2}{E_{y_0}{}^2} + \frac{E_z{}^2}{E_{z_0}{}^2} + 2\frac{E_y E_z}{E_{y_0}} = 0$$

or

$$\{\frac{E_y}{E_{y_0}} + \frac{E_z}{E_{z_0}}\}^2 = 0$$

or

$$E_y = \mp \frac{E_{y_0}}{E_{z_0}} E_z \qquad (11)$$

Thus, resultant relation between E_y and E_z is again a pair coincidence straight lines with a slope of $m = -\frac{E_{y_0}}{E_{z_0}}$ with y-axis. So, the resultant light will be plane polarized light as shown in Fig.5.14 (b).

(B) Production Of Elliptically Polarized Light

(a) If δ is an odd multiple of $\pi/2$ i.e., $\delta = (2n + 1)\pi/2$, the product term in equation (9) reduces to

$$\frac{E_y{}^2}{E_{y_0}{}^2} + \frac{E_z{}^2}{E_{z_0}{}^2} = 1$$

This is the equation of an ellipse, where the minor and major axis i.e., E_{y_0} and E_{z_0} coincides with the coordinate axis. So, if the interfering light waves are polarized and the phase difference between them is odd multiple of $\pi/2$, the resultant light will emerge as elliptically polarized light. Fig.5.15(a) for n = 0 and for n = 1, the rotation of electric vector will be like Fig.5.15 (b).

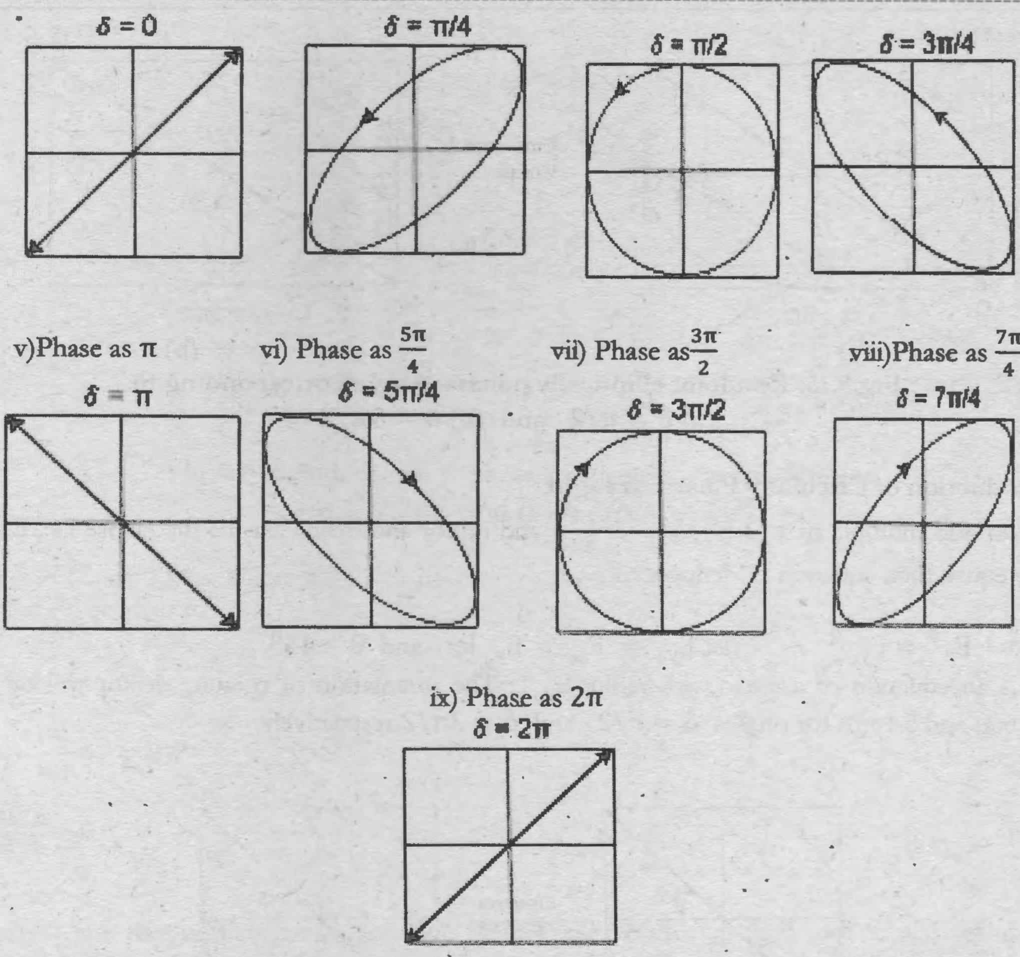


Fig.5.17. States of polarization corresponding to different values of the phase difference from 0 to 2π and θ = 45⁰

5.7 Retardation Plates

Now it has been explained that if light is passed through doubly refracting crystals, there are two refracted rays called O-ray and E-ray. In some crystals called positive crystals, E-rays are slowed down in comparison with O-ray while in other called negative crystals, O-rays are slowed down in comparison to E-rays. Due to this property of these crystals, the specific plates are called as retardation plates. As these crystals are capable to produce two plane polarized light, polarized in perpendicular direction to each other, they can also be used to introduce a required path difference between the two. So, theses plates are widely being used to produce plane, circularly and elliptically polarized light. For this, a doubly refracting uniaxial crystal (e.g., quartz or calcite), whose refracting faces are cut parallel to the direction of optic axis and the light is allowed to incident normally to the optic axis can serve the purpose of retardation plate. The path difference and the phase difference between the two interfering light will be calculated as follows;

Let t be the thickness of the plate in the direction of propagation i.e. perpendicular to the optic axis, μ_0 the index for the O ray; μ_E the index of the plate for the E ray. So, the optical path for the E ray is $\mu_E t$ and for the O ray is $\mu_0 t$ inside the plate.

The path difference for the negative crystals like calcite ($\mu_0 > \mu_E$) is therefore,

$$\Delta = (\mu_0 - \mu_E)t$$

The path difference, however, for the positive crystals like quartz ($\mu_E > \mu_0$) is:

$$\Delta = (\mu_E - \mu_0)t$$

Correspondingly, the phase difference between the two waves for negative crystals and positive crystal respectively, are therefore, expressed as:

$$\delta = \frac{2\pi}{\lambda}(\mu_0 - \mu_E)t$$

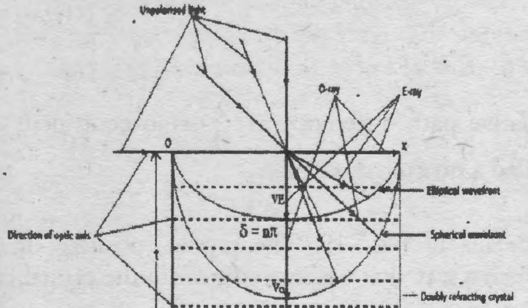
$$\delta = \frac{2\pi}{\lambda}(\mu_E - \mu_0)t$$

There are two commonly used retardation plates:

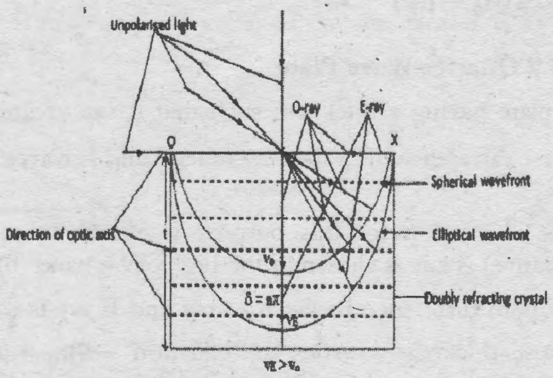
5.7.1 Half-Wave Plate

A plate having a thickness such that it can create a relative path difference of $\frac{\lambda}{2}$ or an equivalent phase difference of π between the O and E waves is called a half-wave plate.

This action is illustrated in more detail in Fig.5.18. The crystal positive or negative is cut as shown in the Fig. 5.18, such that on emerging from the crystal, the path difference between O-ray and E-ray is multiple of $\frac{\lambda}{2}$ or the phase difference is $n\pi$, if the unpolarised light is incident perpendicular to the optic axis. If the light is incident at any angle to the optic axis, we can get the corresponding phase differences as in Fig.5.18(a) and 5.18 (b)



(a)Positive Crystal



(b)Negative Crystal

Fig5.18. Phase difference between O- ray and E-ray

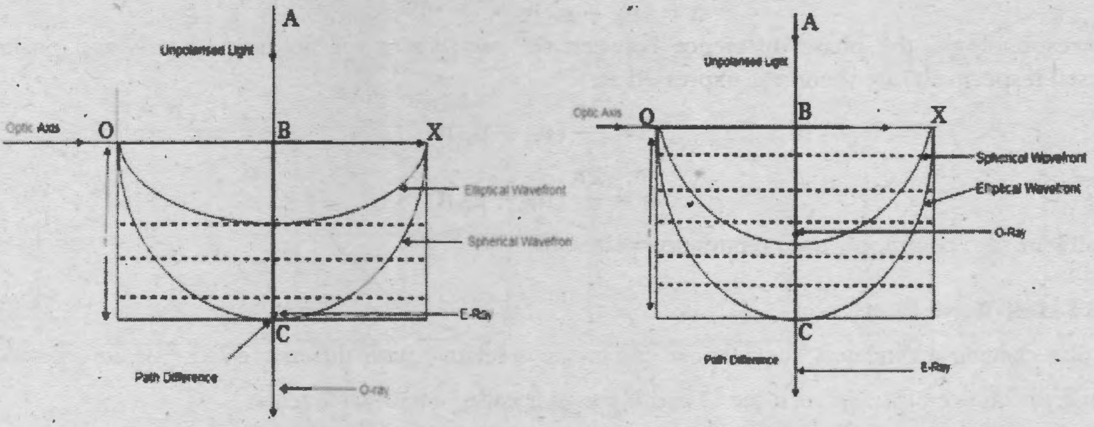
The thickness of a half wave plate will be calculated as;

For the negative crystals like calcite ($\mu_o > \mu_E$) as $v_E > v_0$, therefore:

$$\Delta = (\mu_o - \mu_E)t = n\frac{\lambda}{2}$$

Or $$t = \frac{n\lambda}{2(\mu_o - \mu_E)}$$

Similarly, the path difference for positive crystal can be calculated. The exact diagrams of half wave plates are as Fig.5.19(a) and 5.19(b) and the path difference on the exit point C of the either crystal will be $\Delta = (\mu_o \sim \mu_E)t = \frac{\lambda}{2}$



(a) Positive Plate **(b) Negative Plate**

Fig.5.19. Half Wave Plates

The multiple thicknesses of half wave length are required to produce plane polarized light of positive or negative slopes with the axis otherwise the thickness of a half wave plate to produce a plane polarize light should be:

$$t = \frac{\lambda}{2(\mu_o - \mu_E)}$$

5.7.2 Quarter-Wave Plate

A plate having a thickness such that it can create a relative path difference of $\frac{\lambda}{4}$ or an equivalent phase difference of $\frac{\pi}{2}$ between the O and E waves is called a quarter-wave plate.

The illustration for this purpose is shown in more details in Fig.5.19. The crystal (positive or negative) is cut as shown in the Fig. 5.19 (a) and (b), in such a way that on emerging from the crystal, the path difference between O-ray and E-ray is $\frac{\lambda}{4}$ or the phase difference is $\frac{\pi}{2}$. In general, a plate whose thickness satisfies the equation $t(\mu_o - \mu_E) = (4n + 1)\lambda/4$ behaves like a quarter-wave plate.

The thickness of a quarter wave plate can be well calculated.

The path difference for the positive crystals like quartz ($\mu_E > \mu_o$) is given as:

$$\Delta = (\mu_E - \mu_o)t = (\mu_E - \mu_o)\,(4n + 1)\lambda/4$$

or $t = \frac{(4n+1)\lambda}{4(\mu_o - \mu_E)}$

In particular, the thickness of a quarter wave plate is:

$$t = \frac{\lambda}{4(\mu_o - \mu_E)}$$

The quarter wave plates are widely being used to produce elliptically polarized and circularly polarized light. The symbolic diagrams for the quarter wave plate (positive or negative) will be same as shown in Fig. 5.19 (a) and (b). The only difference is that the thickness is such that on the exit point C of the either crystal, the path difference between O-ray and E-ray will be:

$$\Delta = (\mu_o \sim \mu_E)t = \frac{\lambda}{4}$$

5.8 Analysis of Light

The nature of a beam of light from the polarization aspect can be determined very easily whether it is an unpolarised, polarized, or a mixture etc. through various experiments. We can determine its nature keeping the following different possibilities:

(a) Unpolarised light (b) Plane polarised light (c) Elliptically polarised light (d) Circularly polarised light (e) Mixture of unpolarised and plane polarised (f) Mixture of unpolarised and elliptically polarised light (g) Mixture of unpolarised and circularly polarised light.

If one knows about the method of generation and properties exhibited by each type of the light beam described above, the identification can be done very easily using a Nicol's prism and a quarter wave plate. Following are the steps to test the light:

(I) Light is passed through an analyser (mostly Nicol Prism), rotating about axis as the direction of propagation of light. The intensity of emergent light is then analysed with a photo diode. At the first instance, there may be following three cases;

(i) The intensity does not vary. The inference of this can be concluded as unpolarised light or circularly polarized light as in Fig. 5.20 (i) and (ii)

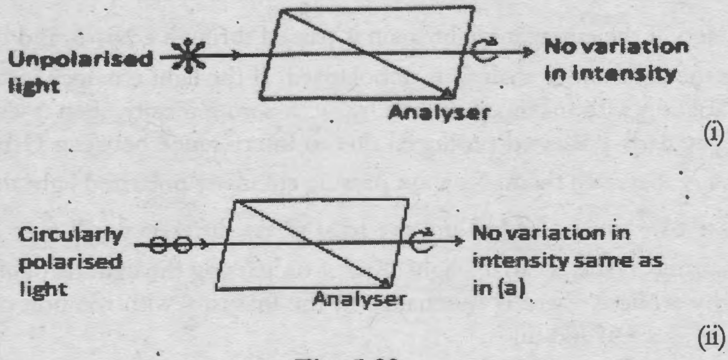


Fig. 5.20.

(ii) During one complete rotation of the Nicol, the intensity varies from maxima to minima two times each. Two positions for maxima are at the ends of one diameter and for minima on other diameter. The diameters are perpendicular to each other. If the minima is of nearly zero intensity, the light is said to be plane polarized light as in Fig. 5.21

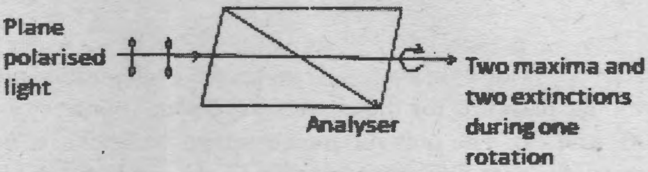
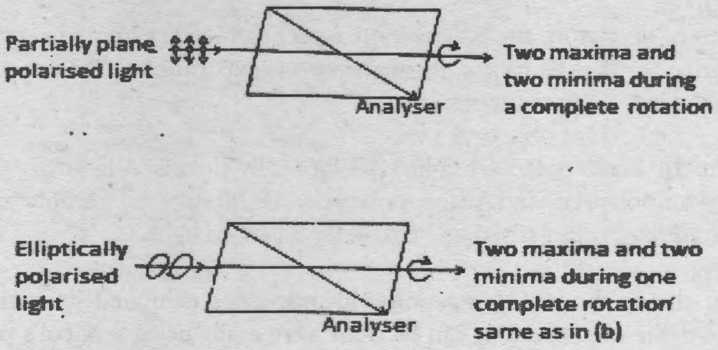


Fig. 5.21.

(iii) If the variation of intensity from maxima to minima shows non-zero minimum intensity, this conclude that light may be partially polarized or elliptically polarized as in Fig.5.22 (i) and (ii).



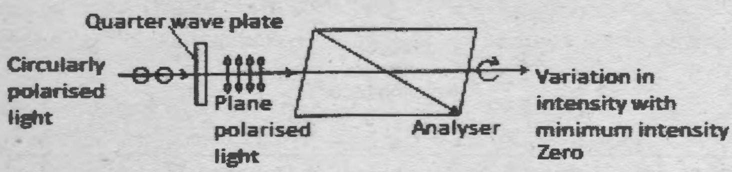
(i)

(ii)

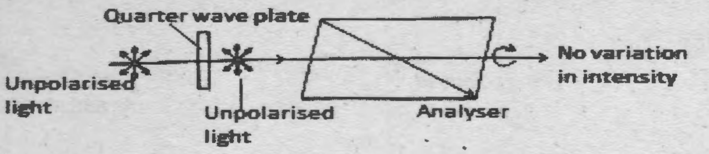
Fig. 5.22.

Now the analysis given in (ii) condition confirms plane polarized light but (i) and (iii) are further to be analysed as follows;

(II) The light is passed through a quarter wave plate, which is capable to produce two plane polarized light, polarized in perpendicular direction and introducing a path difference of $\frac{\lambda}{4}$, or phase difference of $\frac{\pi}{2}$. Now if the emergent light again is passed through a Nicol, and there is no change in the intensity, then the light under analysis is unpolarised. If the light emergent from the quarter wave plate shows the variation with maxima and minima with zero intensity, then it will be concluded that initially light was circularly polarized produced due to interference between O-ray and E-ray with a phase difference of $\frac{\pi}{2}$ between them. Now on passing circularly polarized light through quarter wave plate an extra phase of $\frac{\pi}{2}$ will be added and the total phase difference will now be π, resulting in a condition which satisfies plane polarized light. And if on passing the light through quarter wave plate and on analyzing by a Nicol, there is no change in the intensity with rotation of Nicol, the light is unpolarised as in Fig. 5.23 (i) and (ii)



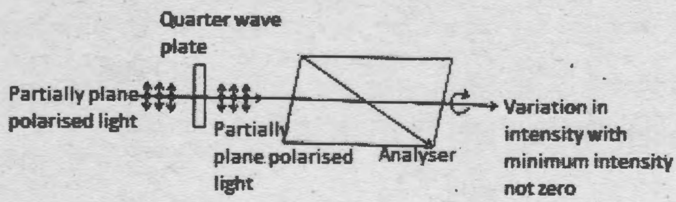
(i)



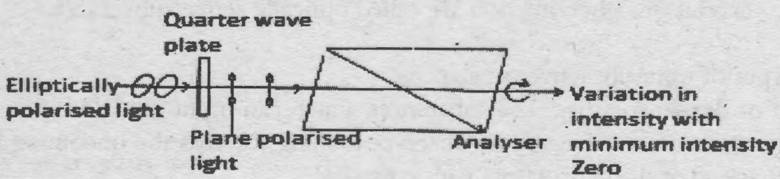
(ii)

Fig. 5.23.

(III) Now to distinguish partially plane polarized and elliptically polarized light, first the light is passed through a quarter wave plate to introduce extra phase difference of $\frac{\pi}{2}$. If initially the light was elliptically polarized, it will convert plane polarized and on analyzing through a Nicol, variation of intensity will be maxima to minima with zero intensity. And if initially the light was partially plane polarized, on introducing extra phase of $\frac{\pi}{2}$ the light will convert to elliptically polarised light showing variation in the intensity with non-zero intensity with Nicols in Fig. 5.24 (i) and (ii)



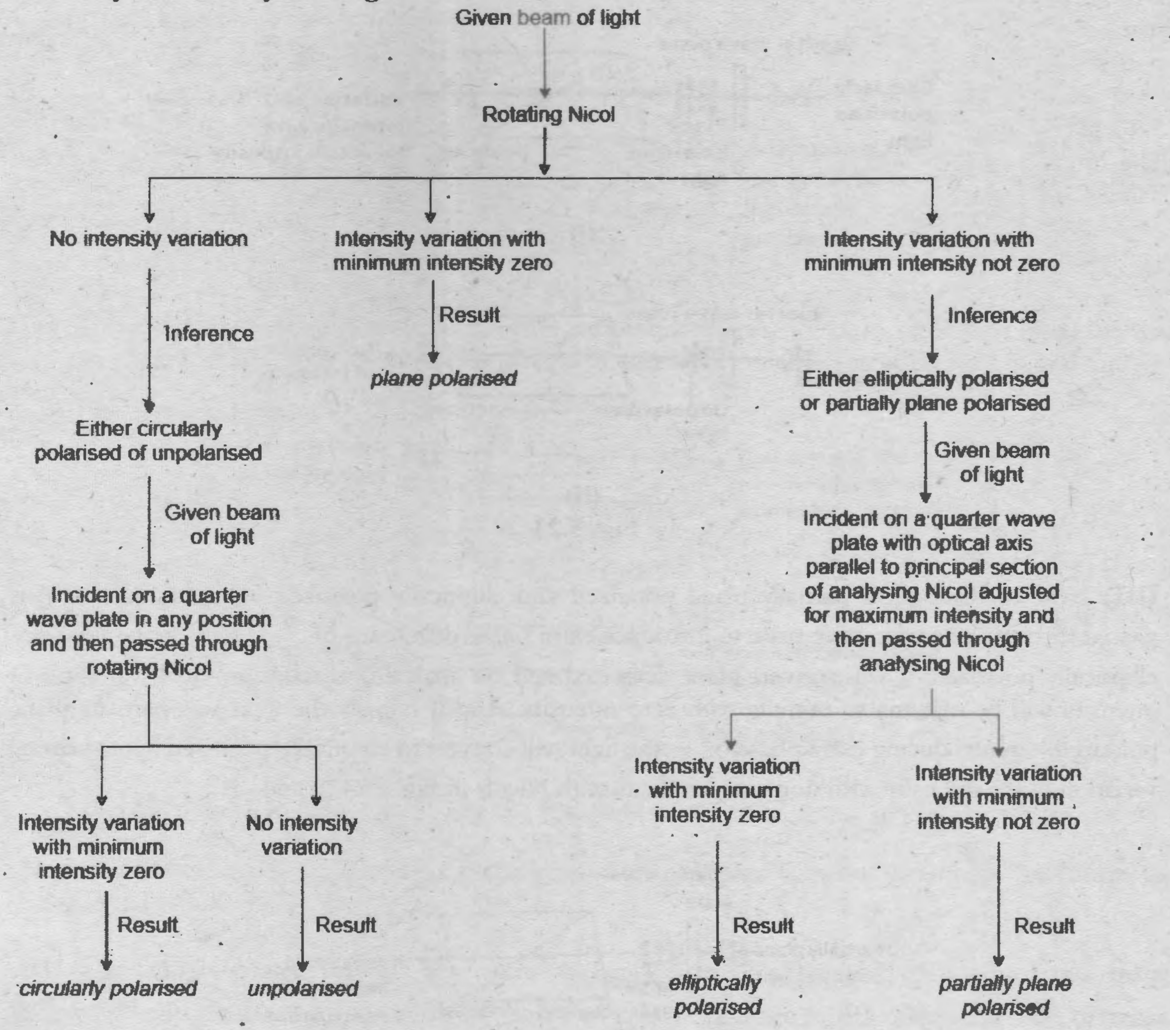
(i)



(ii)

Fig. 5.24.

Summary of the analysis of light



5.9 Optical Activity: Rotatory Polarisation

The property or phenomenon of rotation of the plane of polarisation about the direction of propagation of light by certain substances is called optical rotation or optical activity and the substances which exhibit this phenomenon are called optically active substances.

There are two types of optically active substances:

(i) Right-handed or dextro-rotatory: The substances which can rotate the plane of polarisation in the clockwise direction with respect to an observer looking towards the oncoming beam of light are called right-handed or dextro-rotatory e.g., cane sugar

(ii) Left-handed or laevo-rotatory: The substances which can rotate the plane of polarisation in the anti-clockwise direction with respect to an observer looking towards the oncoming beam of light are called left-handed or laevorotatory e.g., fruit sugar.

Optically active substances thus have an asymmetric molecular structure such that their molecules may exist either in left or right-handed form such that the left-handed substances rotate the plane of polarization to the left i.e., anticlockwise and right-handed to the right i.e., clockwise.

5.9.1 Biot's Laws for Rotatory Polarisation

The phenomenon of optical rotation was first studied in detail by Biot, in 1815, with both right-handed and left-handed crystals. His experimentally determined laws are as follows:

1. The angle of rotation (θ) of the plane of vibration, for any given wavelength, is directly proportional to the length (l) of optically active substance traversed by the incident light
$$\theta \propto l$$
2. In the case of solutions, the amount of rotation (θ) is directly proportional to the concentration (c) of optically active substance in the solution.
$$\theta \propto c$$
3. The angle of rotation is approximately inversely proportional to the square of the wavelength for a given thickness of an optically active substance
$$\theta \propto \frac{1}{\lambda^2}$$
Angle of rotation (θ) is least for red and greatest for violet. White light passes through optically active substances, different colours may undergo different amount of rotation, also known as rotatory dispersion.

4. The rotation produced by a number of optically active substances is equal to the algebraic sum of the individual rotations produced by these substances separately. Mathematically, we can express the resultant optical rotation θ due to a number of substances arranged one after another as:
$$\theta = \theta_1 - \theta_2 + \theta_3 - \theta_4 \ldots\ldots\ldots\ldots \theta_n$$
Where the substances are alternatively right-handed and left-handed varieties. The rotation in the clockwise direction is taken as positive while that in anticlockwise direction is taken as negative. A solution containing equal number of right- and left-handed molecules is optically inactive.
5. The amount of rotation also depends on the nature of the substance and it decreases with temperature of the substance.

This phenomenon can be utilized in finding the percentage of the optically-active substance, present in a given base solution. For example, this phenomenon is being used in the sugar industry for measuring the percentage of sugar in sugar-cane juice. Also, by measuring the angle of rotation of the plane of polarisation, the amount of sugar present in the urine of a diabetic patient may be readily determined.

5.9.2 Optical Rotation: Fresnel's Theory

The phenomenon of rotatory polarization was first explained by Fresnel in the optically active crystals like quartz. His theory is based on the principle that any rectilinear simple harmonic motion is equivalent to the superposition of equal and oppositely described circular vibrations one clockwise and the other anti-clockwise- rotating with equal frequency.

Assumptions of Fresnel's Theory

(i) When a plane polarised beam enters a crystal along the optic axis, it decomposes up into two oppositely directed circularly polarised beams; one clockwise and the other anti-clockwise. These are propagated along the optic axis without any change.

(ii) The two circularly polarised beams travel with the same velocity for an optically in-active substance and as a result both the beams arrive at a given point along their path simultaneously. The resultant therefore, will be a simple harmonic motion in the path of the incident vibration. The vibrations therefore will be in the same plane for a plane polarized light which is incident along the axis.

For an optically active substance, the two beams travel with different velocities. In a dextro-rotatory substance, the velocity of right-handed circularly polarised light is greater than that of left-handed circularly polarised light; while in laevorotatory substance, the velocity of left-handed circularly polarised light is greater than that of right-handed. Thus, a phase difference is being introduced between the two circularly polarized light beams while traversing through the optically active substances.

(iii) When the two circularly polarised beams emerge from the optically active substance, they recombine to form a plane polarised beam. Their plane of polarisation is rotated with respect to that of the incident light through a certain angle which depends on the phase difference between the two beams.

Consider a plane-polarised beam incident normally on a quartz crystal cut with its face perpendicular to the optic axis. Let the vibration in the incident light be represented by:

$$y = a \sin \omega t \tag{12}$$

Where a is the amplitude.

According to Fresnel's assumptions, the incident beam breaks up into two oppositely directed circularly polarised beams as shown in Fig.5.25 (a) when it enters the quartz plate.

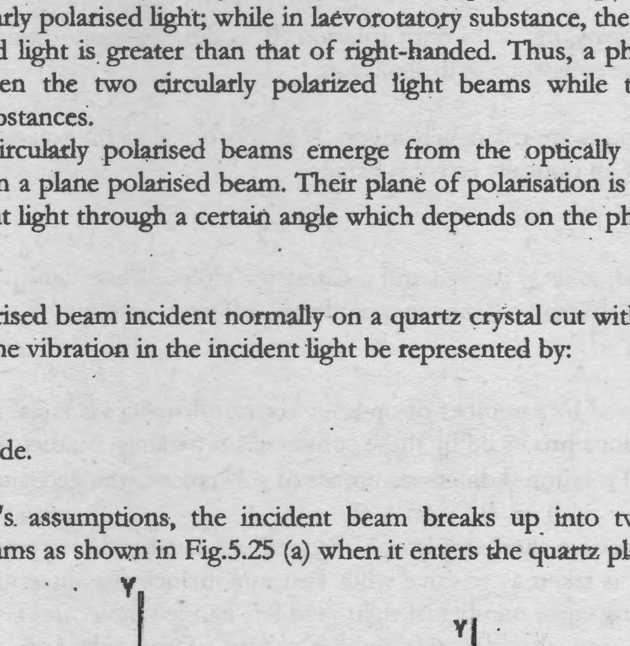


Fig. 5.25.

These circular vibrations are represented by the equations:

$$x_1 = \frac{a}{2}\cos\omega t, \qquad y_1 = \frac{a}{2}\sin\omega t$$

And $x_2 = -\frac{a}{2}\cos\omega t$, $y_2 = \frac{a}{2}\sin\omega t$

Since the two circularly polarized light beams traverse through the optically active substance with different velocities, therefore, a phase difference is being introduced when they emerge from the quartz plate, Fig. 5.25 (b).

The emergent circular vibrations are represented by the equations:

$x_1 = \frac{a}{2}\cos\omega t$, $y_1 = \frac{a}{2}\sin\omega t$

$x_2 = -\frac{a}{2}\cos(\omega t + \delta)$; $y_2 = \frac{a}{2}\sin(\omega t + \delta)$

The resultant vibration is:

$$X = x_1 + x_2 = \frac{a}{2}\cos\omega t - \frac{a}{2}\cos(\omega t + \delta) = a\sin\frac{\delta}{2}\sin(\omega t + \frac{\delta}{2}) \quad (12)$$

and $Y = y_1 + y_2 = \frac{a}{2}\sin\omega t + \frac{a}{2}\sin(\omega t + \delta) = a\cos\frac{\delta}{2}\sin(\omega t + \frac{\delta}{2})$ (13)

where X and Y are two perpendicular vibrations in the same phase.

Divide (12) by (13)

$$\frac{X}{Y} = \tan\frac{\delta}{2} \quad (14)$$

This shows the equation of a straight line inclined at an angle $\frac{\delta}{2}$ with Y axis. We can say that the plane polarized light emerging out from the quartz plate results in the vibrations inclined at an angle $\frac{\delta}{2}$ to the Y axis i.e to the vibrations in the incident light. Since the phase difference between the two circularly polarized vibrations is given by δ, the angle of rotation θ of the plane of polarization is half the difference between the circularly polarized vibrations.

If the refractive indices of quartz along optic axis of left-handed (anti-clockwise) and right-handed (clockwise) circularly polarised light be μ_L and μ_R respectively and d is the thickness of the quartz plate, the path difference between two beams is given as:

$$= (\mu_L - \mu_R)d$$

The phase difference, $\delta = \frac{2\pi}{\lambda}(\mu_L - \mu_R)d$

Hence the rotation of the plane of polarisation is given by,

$$\theta = \frac{\delta}{2} = \frac{\pi}{\lambda}(\mu_L - \mu_R)d \quad (15)$$

If v_L and v_R are the velocities of the left-handed and right-handed circularly polarised light, we have

$\mu_L = \frac{c}{v_L}$ and $\mu_R = \frac{c}{v_R}$

where c is the velocity of light.

Substitute the values in (15) we get,

$$\theta = \frac{\pi d}{\lambda}\left(\frac{c}{v_L} - \frac{c}{v_R}\right) = \frac{\pi d}{T}\left(\frac{1}{v_L} - \frac{1}{v_R}\right) \quad \text{as } \frac{c}{\lambda} = \frac{1}{T} \tag{16}$$

Since, $v_R > v_L$, therefore $\mu_L > \mu_R$.

Also, $\theta \propto d$, this implies that plane polarized light continues to travel in the right-handed crystal along its optic axis whenever the plane of vibration rotates in the clockwise direction. For a left-handed crystal, it would be exactly opposite. When $\theta = 0$, $\mu_L = \mu_R$ and calcite will not show any optical rotation this implies that the plane polarized light is propagated along the optic axis of such crystals with the direction of vibrations unaffected.

Experimental Verification of Fresnel's Theory

Fresnel measured experimentally that a plane polarised light resolves into two circularly polarised vibrations which travel with different velocities when plane polarized light enters inside the crystal.

If the assumptions of Fresnel's experimental arrangement are correct, then a prism of quartz can separate out the right-handed (R) and left-handed (L) circular vibrations (Fig. 5.26).

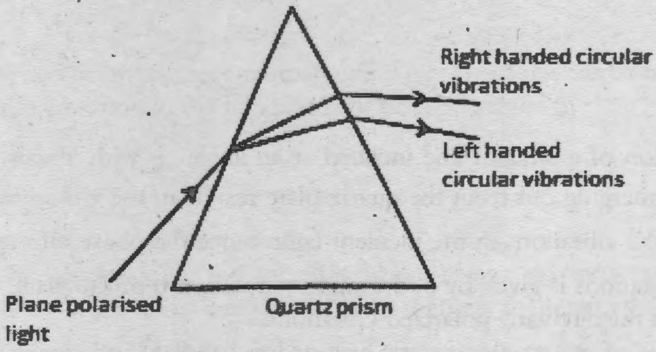


Fig.5.26.

In Fresnel's experimental arrangement, a large number of right-handed (R) and left-handed (L) quartz prisms alternatively are combined as shown in (Fig.5.27).

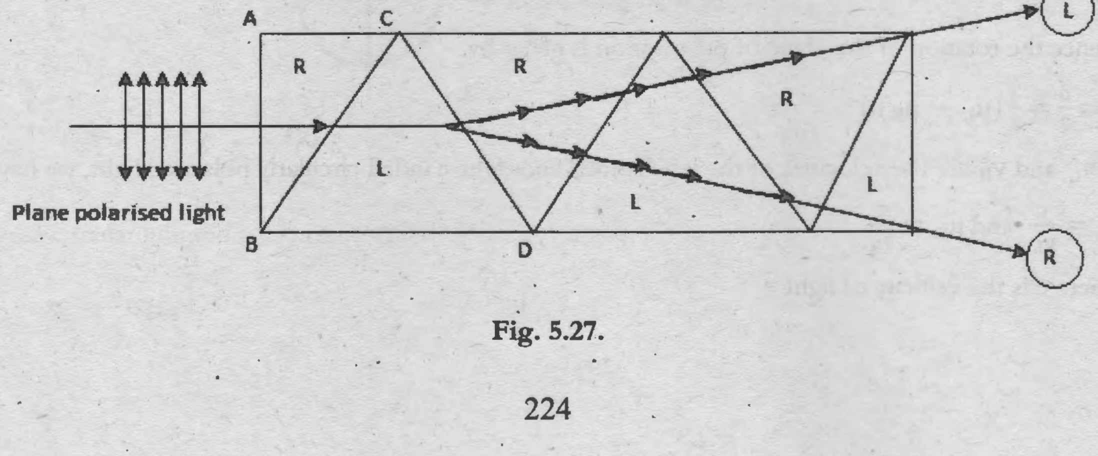


Fig. 5.27.

Each prism, has its optic axis parallel to its base. Let us verify the theory according to which a beam of plane polarised light is allowed to fall normally on the face AB of first prism. If Fresnel's hypothesis is correct, this plane polarised beam splits up into two circularly left-handed and right-handed polarised beams which travel through the first prism along the same direction; but with different speeds. In case of a right-handed quartz prism, right-handed circularly polarised beam travels faster than left-handed beam. As soon as beam is incident on the oblique surface BC of second prism (L), right-handed beam becomes slower than left-handed and vice a versa. Hence, the second prism (L) is optically denser medium for right-handed beam and rarer for left-handed beam. Thus, in second prism right-handed beam bends towards the normal (i.e., downward), while the left-handed beam bends away from the normal (i.e., upward). At the surface CD, the velocities are again interchanged, so that right-handed beam bends away from the normal on CD, i.e., downward and left-handed beam bends upward. Thus, at successive surfaces of the prism, the right-handed beam bends downward and left-handed beam bends upward i.e., the angular separation between two circularly polarised beams increases continuously. Separation between two circularly polarised beams can easily be detected by using a suitable number of prisms. When the beams emerging from the system are analysed by a Nicol prism and a quarter wave plate, it is found that both the beams are circularly polarised in opposite directions. Thus, Fresnel's theory is experimentally verified.

5.10 Specific Rotation

The dependence of the rotation produced by an optically active substance dissolved in a non-active solvent upon the concentration of solution has led to a large number of applications e.g., in the estimation and analysis of, sugar concentrations and many other optically active chemical solutions. For comparison of optical activity, therefore a unit is adopted which is known as specific rotation. The amount of the rotation or the specific rotatory power depends upon:

(i) the length of the medium,
(ii) concentration of the solution or the density of optically active substance in the solvent,
(iii) wavelength of light and
(iv) temperature.

The specific rotation of a solution at a given temperature and for a given wavelength of light is thus defined as the rotation (in degrees) produced by 1 decimetre (10 cm) length of the solution when its concentration is 1 g per cubic centimetre i.e.,

$$\alpha_{\lambda}^{t} = \frac{\theta}{lc} \qquad (17)$$

Where α represents specific rotation at temperature 't' for wavelength 'λ', l is the length of the solution in decimetres and c is the concentration of the optically active substance in g. per c.c. in the solution. The unit of specific rotation is $\text{deg}\ (\text{decimetre})^{-1}(\text{g/cm}^3)^{-1}$. The specific rotation for the dextro substances is written positive and for laevo-rotatory is negative.

The specific rotation depends on the temperature and wavelength of incident light. In other words, optical rotation is more for shorter wavelengths and vice a versa.

Rotatory Dispersion

Different amount of rotation for the lights of different wavelength is known as rotatory dispersion. This is explained as follows.

As we know, the angle of rotation of the plane of polarisation depends upon,

(i) the thickness of the medium traversed, i.e., path length,
(ii) the concentration of the solution,
(iii) the wavelength of light and
(iv) temperature

For a solution of given concentration, given path length and at a given temperature, the angle of rotation of plane of polarisation varies approximately as $1/\lambda^2$. Therefore, if white light is passed through an optically active solution, the angle of rotation of the plane of polarisation for different wavelengths will be different, being minimum for red and maximum for violet. This phenomenon is called rotatory dispersion.

5.11 Polarimeters - Optical Activity Measuring Device

The device which measures the angle through which an optically active substance rotates the plane of polarisation of a plane polarised beam is called a polarimeter.

The earlier simple polarimeters consisted of two Nicols (polarizer and analyser) capable of rotation about a common axis and kept a distance apart. This was first deployed by Milcherlich. The attempt of accurately finding the angle of rotation though failed owing to the difficulty in accurately examining one position of the analyser for complete extinction of the transmitted polarized light. This is due to the fact that the analyser could be rotated through an appreciable angle from its setting for complete extinction without allowing any appreciable amount of light to pass through it. Thus, the field of view remains totally dark not for a single, position but for a considerable range of the rotation of the analyser.

To overcome this, polarimeters were designed so as to increase the sensitivity of the pair of crossed Nicols, on the principle of half shade. The use of half shade divides the field of view in two halves. As a result, the analyser can be set accurately for the equality of brightness of the two halves. It should always be remembered that the eye is a better judge of equality of brightness than either the point of complete extinction or maximum brightness of the whole field of view.

There are two types of polarimeters in use: (a) Laurent's half-shade polarimeter, (b) Bi-quartz polarimeter.

5.11.1 Laurent's Half-Shade Polarimeter

Construction

The introduction of half shade between the two Nicols could overcome the drawbacks of the simple polarimeter used before. This was named as Laurent half shade as in Fig. 5.28, named after the scientist Laurent.

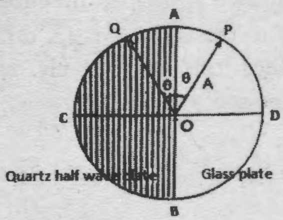


Fig.5.28. Laurent half shade

This comprise of two semi-circular pieces: one of quartz cut parallel to the optic axis and the other of glass. The thickness of the quartz plate is such that it introduces a path difference $\lambda/2$ or a phase difference between the ordinary and the extraordinary vibrations. As seen from the Fig. 5.28, the semi-circular piece ACB is a half wave plate made of quartz and ADB of glass joined together along AB so as to form a composite circular plate CADB. The thickness of glass half is also adjusted such that it transmits the same amount of light as the quartz half does.

Behind the polarizer, we place this circular plate known as half shade such that the quartz and glass halves both cover a semi-circle each of the field of view. Its action can be understood as:

Action of Half Shade

Let the plane of vibration of the light incident on the composite glass- quartz plate be parallel to OP inclined at an angle θ to the optic axis AB of the quartz half. It is seen that the vibrations of light transmitted through the glass half will remain in the same plane as in Fig.5.29. On the other hand, there occurs a change in the quartz half as in Fig. 5.30.

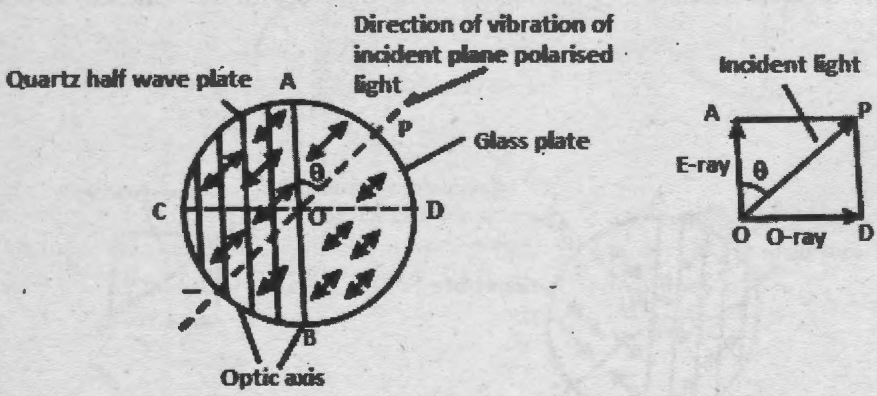


Fig. 5.29.

Let us consider the vibrations incident upon the first face of the half shade is given by:

$$P_i = A \sin\omega t$$

where A = amplitude and subscript 'i'= for incident wave

We can resolve the amplitude A along the two principal directions of the quartz half i.e., along the optic axis (Y direction) and perpendicular to the optic axis (X direction). $A\sin\theta$ and $A\cos\theta$ be then the components of A along X and Y directions respectively. Hence the incident plane polarized vibrations contained in the plane OP will be written as:

$x_i = A\sin\theta\sin\omega t$ [O wave]
$y_i = A\cos\theta\sin\omega t$ [E wave]

These two mutually perpendicular vibrations travel with unequal speeds along the same direction perpendicular to both the X and Y directions defined above. As a result, the two vibrations emerging out from the quartz half will suffer a phase difference as the quartz half is essentially a half wave plate. We then have:

$$x_e = A\sin\theta\,\sin(\omega t + \pi)$$
$$= -A\sin\theta\,\sin\omega t$$
$$y_e = A\cos\theta\,\sin\omega t$$

Where subscript 'e' stands for the emergent beam. The resultant vibration will be given by:

$$R = \sqrt{x_e^2 + y_e^2} = A\sin\omega t$$

and the angle ϕ which the plane of R makes with the axis of X is given by:

$$\tan\phi = \frac{y_e}{x_e} = -\cot\theta$$
$$\phi = \left(\frac{\pi}{2} + \theta\right)$$

Thus, the vibration plane of the light emergent from the glass half is OP while that of the light emerging from the quartz half is OQ making the same angle θ with the optic axis as does OP as in Fig. 5.30

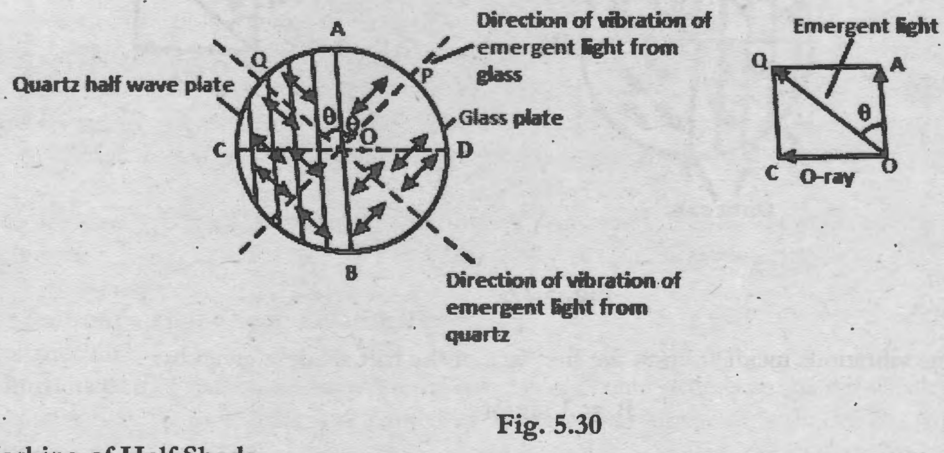


Fig. 5.30

Working of Half Shade

This can be explained with the help of a diagram as shown in Fig.5.31.

i) If the principal plane of the analyser is parallel to COD, the components OQ′ and OP′ of the vibrations emerging out of the two halve will be equal and thus they will appear equally bright [Fig.5.31 (a)].

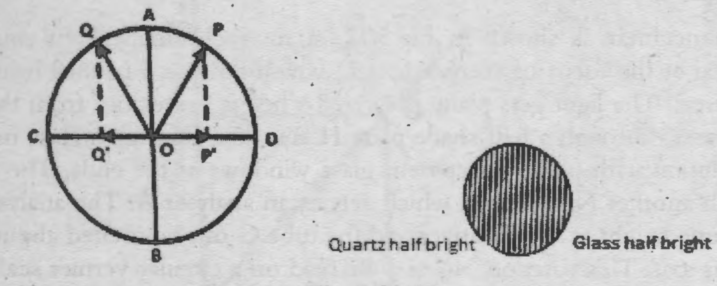


Fig.5.31.(a)

i) If the principal plane of the analyser is rotated by a small angle in the clockwise direction Fig.5.31 (b) the vibrations from the glass half will be completely cut off (provided OP is normal to C′D′) while a component OQ′ of the vibrations emerging out of the quartz half will be transmitted. In general, the glass half will be less bright than quartz half.

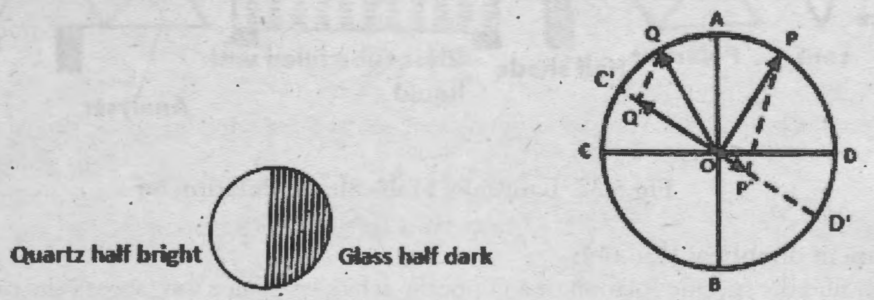


Fig.5.31.(b)

iii) If on the other hand the analyser is turned through a small angle in the anticlockwise direction Fig.5.31(c) the component of the vibration emerging from the glass half will decrease and if OQ is normal to C′D′ the quartz will be dark and the glass half will be bright.

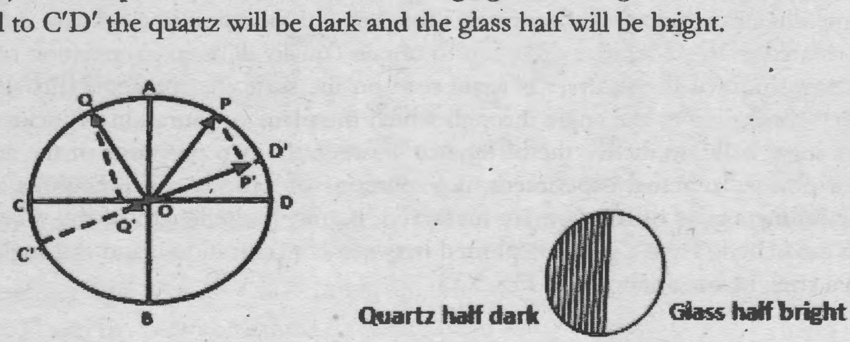


Fig.5.31.(c)

Note: The half shade serves the purpose of dividing the field of view in two halves. The device is so accurate that if the principal plane of the analyser is rotated through even a small angle with respect to COD, a marked change in the intensity of the two halves is observed.

Experimental Set up

The experimental arrangement is shown in Fig.5.32. A monochromatic light source S (usually a sodium lamp) is placed at the focus of convex lens L, which renders a parallel beam that falls on a Nicol Prism P (polariser). The light gets plane polarized when it passes out from the polarizer. This polarized light then passes through a half shade plate H and then through a glass tube G containing an optically active solution with plane transparent glass windows at the ends. The transmitted light further passes through another Nicol Prism which acts as an analyser A. The analyser A is mounted coaxially and at the same height as the polarizer and the tube G can be rotated about the direction of propagation of light as axis. This rotation can now be read on a circular vernier scale graduated with degrees. The light emerging out can be viewed using a telescope T focused on H.

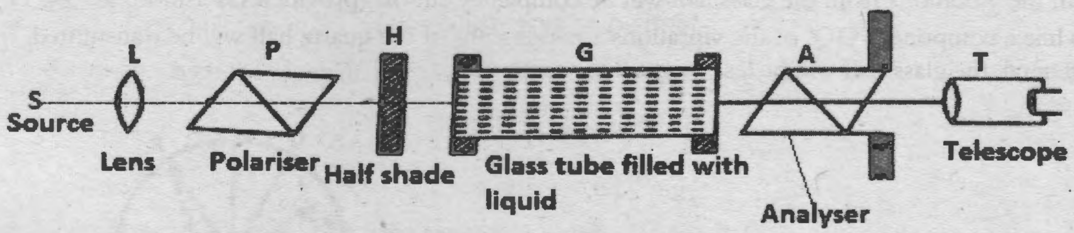


Fig.5.32. Laurent's Half- Shade Polarimeter

Measurement of optical Rotation

In order to find the specific rotation of an optically active substance (say, sugar), the tube is first filled with water and the analyser is adjusted to obtain the condition of equal brightness of the two halves of the field of view. The position of the analyser is read on the scale. Now the sugar solution of known concentration is filled in the tube G (without any air gap). The solution rotates the planes of vibration OP and OE (i.e., of light emerging from glass half and the quartz half) through the same angle in the same direction. Sugar solution rotates the plane of vibration in clockwise direction. Now the analyser is rotated in the clockwise direction to obtain equally-illuminated position of the field of view again. This position of the analyser is again read on the scale. As the angle through which the analyser has been rotated gives the angle through which the plane of vibration of incident beam has been rotated by sugar solution, hence the difference between the two positions of the analyser gives the angle of rotation θ. In actual experiment, the solutions of various concentrations are prepared and the corresponding angles of rotation are measured. Knowing the length of the tube the specific rotation can be calculated. Then a graph is plotted between concentration c and the angle of rotation θ. The graph is a straight line as shown in Fig. 5.33

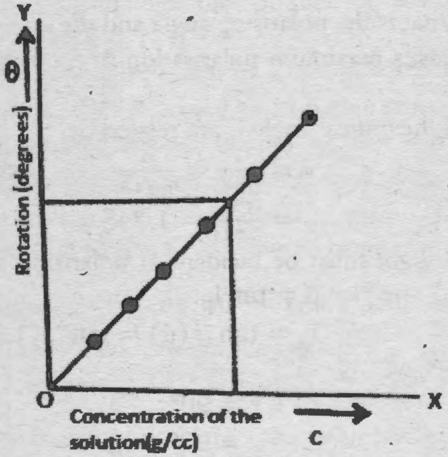


Fig.5.33. Graph of θ vs c

Knowing the length of the tube, the specific rotation can also be calculated. From the graph, from a set of θ and c, slope can be found and the specific rotation of sugar is then calculated from the following relation

$$\alpha_\lambda^t = \frac{10\theta}{lc}$$

Where l is the length of the tube in cm., c is the concentration of the solution in g./c.c. and θ is the rotation in degrees.

*******Solved Examples*******

Based On Brewster Law

Ex. 1. A ray of light is incident on the surface of a glass plate of refractive index 1.732 at the polarising angle. Calculate the angle of refraction of the ray.

Sol. According to Brewster's law $\mu = \tan i_p$

Here $\mu = 1.732$ $\therefore$ $1.732 = \tan i_p$

or $i_p = \tan^{-1} 1.732 = \tan^{-1}\sqrt{3} = 60°.$

If r is the angle of refraction, we have

$$r + i_p = 90°$$

$$r = 90° - i_p = 90° - 60° = 30°.$$

Ex. 2. A beam of light travelling in water strikes a glass plate which is also immersed in water. When the angle of incidence is 51°, the reflected beam is found to be plane polarised. Calculate the refractive index of glass.

Sol. By Brewster's law, $\mu = \tan i_p$

Here $i_p = 51°$ and the beam of light is travelling from water to glass.

$$\mu_g^w = \tan i_p = \tan 51° = 1.235.$$

This is the refractive index of glass with respect to water is 1.235.

Ex.3. At a certain temperature the critical angle of incidence of water for total internal reflection is 48° for a certain wavelength. What is the polarising angle and the angle of refraction for light incident on the water at an angle that gives maximum polarisation of the reflected light? (Given $\sin 48° = 0.7431$).

Sol. The refractive index μ and the critical angle C are related by

$$\mu = \frac{1}{\sin C} = \frac{1}{\sin 48°} \qquad [\text{since } C = 48° \text{ (given)}]$$

$$= \frac{1}{0.731} = 1.345$$

For maximum polarisation, the light must be incident at polarising angle. From Brewster's law we have $\mu = \tan i_p$.

∴ Polarising angle, $i_p = \tan^{-1}(\mu) = \tan^{-1}(1.345) = 53°22'$

If r is the angle of refraction we have

$$i_p + r = 90°$$

$$r = 90° - i_p$$

$$r = 90° - 53°22' = 36°38'.$$

Based On Malus law

Ex.4. Two polarising plates have polarising directions parallel so as to transmit maximum intensity of light. Through what angle must either plate be turned if the intensity of the transmitted beam is to drop by one-third.

Sol. According to Malus, we have $I = I_0\cos^2\theta$.

Here $I = \frac{I_0}{3}$ $\quad\therefore \frac{I_0}{3} = I_0\cos^2\theta$

$\cos^2\theta = \frac{1}{3}$ or $\cos\theta = \pm\frac{1}{\sqrt{3}}$

∴ $\theta = \cos^{-1} \pm (0.5773) = \pm 55°18'$ or $\pm 145°18'$.

Ex.5. A beam of light is passed through two Nicols in series. In a particular setting maximum light is passed by the system and it is 500 units. If one of the Nicols is now rotated by 20°, calculate the intensity of transmitted light.

Sol. We have $I = I_0\cos^2\theta$.

Here $I_0 = 500$ units, $\theta = 20°$

$$I = 500 \times \cos^2 20° = 500 \times (0.9397)^2$$

$$= 500 \times 0.883 = 442 \text{ units.}$$

Ex. 6. Two Nicols are oriented with their principal planes making an angle of 30°. What percentage of incident unpolarised light will pass through the system?

Sol. Let the intensity of the incident unpolarised light be $2I_0$. On entering the first nicol it is broken up into O and E-components each of intensity I_0. The O component is totally reflected and only E component of intensity of I_0 is transmitted through this (first) nicol. According to Malus law, the intensity of finally transmitted beam, is given by

$$I = I_0\cos^2\theta = I_0\cos^2 30° = I_0\left(\frac{\sqrt{3}}{2}\right)^2 = \frac{3}{4}I_0$$

∴ Percentage of incident unpolarised light transmitted through the system

$$= \frac{I}{2I_0} \times 100 = \frac{\left(\frac{3}{4}I_0\right)}{2I_0} \times 100 = \frac{3}{8} \times 100 = 37.5\%.$$

Ex.7. Two nicols are first crossed and then one of them is rotated through 60°. Calculate the percentage of incident light transmitted.

Sol. Let the intensity of incident unpolarised light be $2I_0$. On entering the first nicol it is broken up into O and E-components each of intensity I_0. The O-component is totally reflected and only E-component of intensity I_0 is transmitted through the first nicol.

When the nicols are crossed, the angle between their principal planes is 90°. Now if one of them is rotated through 60°, the angle between their planes of transmission is either 30° or 150°.

According to Malus law, the intensity of finally transmitted beam is given by

$$I = I_0\cos^2\theta = I_0\cos^2 30° = I_0\cos^2 150° = I_0\left(\frac{\sqrt{3}}{2}\right)^2 = \frac{3}{4}I_0$$

∴ Percentage of incident light transmitted through the system

$$= \frac{I}{2I_0} \times 100 = \frac{3I_0/4}{2I_0} \times 100 = \frac{300}{8} = 37.5\%.$$

Ex.8. Two light sources are observed, one after the other with two nicol prisms mounted one behind the other as polariser and analyser. The intensity of beams transmitted by the prisms are observed equal from two sources when the angle between the principal section of the nicols are 45° and 70° respectively. Calculate the relative intensities of the two sources.

Sol. Let the intensities of the given sources be $2I_0$ and $2I'_0$ respectively. On entering the first nicol (polariser), in each case the beams get broken up into O and E-components each of intensity I_0 in first case and I_0' in second case. The O-component is totally reflected and only E-component of intensity I_0 in first case and I_0' in second case is transmitted through the first nicol. According to Malus law the intensities of finally transmitted beams are given by

$$I = I_0\cos^2\theta \text{ for first source}$$

$$I' = I_0'\cos^2\theta' \text{ for second source}$$

But

$$I = I' \text{ (given), therefore}$$

$$I_0\cos^2\theta = I_0'\cos^2\theta'$$

i.e.

$$\frac{I_0}{I_0'} = \frac{\cos^2\theta'}{\cos^2\theta}$$

Hence, the relative intensities of two sources

$$\frac{2I_0}{2I_0'} = \frac{\cos^2\theta'}{\cos^2\theta} = \frac{\cos^2 70°}{\cos^2 45°} = \frac{(0.3420)^2}{(1/\sqrt{2})^2} = 2 \times (0.3420)^2 = 0.2339.$$

Based On Retardation Plates

Ex.9. Calculate the thickness of a quarter wave plate of quartz for sodium light of wavelength 5893 Å. The refractive indices of quartz for E-ray and O-ray are equal to 1.5533 and 1.5442 respectively.

Sol. The thickness of the quarter wave plate of quartz is given by

$$t = \frac{\lambda}{4(\mu_E - \mu_O)}.$$

Here $\lambda = 5893$ Å$= 5893 \times 10^{-8}$cm, $\mu_E = 1.5533$, $\mu_O = 1.5442$

$$t = \frac{5893\times10^{-8}}{4\times(1.5533-1.5442)}$$

$$= \frac{5893\times10^{-8}}{4\times0.0091} = 1.62 \times 10^{-3} \text{ cm}.$$

Ex.10. Calculate the thickness of a calcite plate which would convert plane polarised light into circularly polarised light. The principal refractive indices are $\mu_0 = 1.658$ and $\mu_E = 1.486$ at the wavelength of light used, 5893 Å.

Sol. The plane polarised light will be converted into circularly polarised light if the thickness of the plate introduces a phase difference of $\pi/2$ or an odd multiple of $\pi/2$, i.e., a path difference of $\pi/4$ or an odd multiple of $\pi/4$ between ordinary and extraordinary rays. Hence the thickness 't' of the calcite plate is given by

$$(\mu_0 - \mu_E)t = \frac{(2n-1)\lambda}{4}, n = 1,2,3, ...$$

$$t = \frac{(2n-1)\lambda}{4(\mu_0-\mu_E)}$$

Hence $\lambda = 5890$ Å $= 5890 \times 10^{-8}$cm, $\mu_E = 1.486, \mu_0 = 1.658$

$$t = \frac{(2n-1)\times 5890\times 10^{-8}}{4(1.658-1.486)}$$

$$= \frac{(2n-1)\times 5890\times 10^{-8}}{4\times 0.172}$$

$$= 8.56 \times 10^{-5}(2n-1)\text{cm}, \quad n = 1,2,3 ....$$

∴ Thickness of the calcite plate is given by

$$t = 8.56 \times 10^{-5} \times 18.56 \times 10^{-5} \times 38.56 \times 10^{-5} \times 5 ...$$

$$= 8.56 \times 10^{-5}\text{cm}, 25.68 \times 10^{-5}\text{cm}, 42.80 \times 10^{-5}\text{cm}, ...$$

The least thickness of the plate $= 8.56 \times 10^{-5}$cm.

Ex.11. A given calcite plate behaves as a half wave plate for a particular wavelength λ. Assuming variation of refractive index with λ to be negligible, how would the above plate behave for another light of wavelength 2λ?

Sol. The thickness of a half wave plate of calcite crystal for wavelength λ is given by

$$t = \frac{\lambda}{2(\mu_0-\mu_E)} = \frac{2\lambda}{4(\mu_0-\mu_E)} = \frac{\lambda'}{4(\mu_0-\mu_E)} \quad \text{where } \lambda' = 2\lambda$$

Obviously, the half wave plate for λ will behave as a quarter wave plate for $\lambda'(=2\lambda)$ provided the variation of refractive index with wavelength is negligible.

Ex.12. The faces of a quartz plate are parallel to the optic axis of the crystal (a) What is the thinnest possible plate that would serve to put the ordinary and extraordinary rays of $\lambda = 5890$ Å a half wave apart on their exit? (b) What multiples of this thickness would give the same result? The indices of refraction of quartz are $\mu_E = 1.553$ and $\mu_0 = 1.544$.

Sol. (a) The thinnest possible plate that will produce a path difference of $\lambda/2$ (half wave), is a half wave plate. The thickness of a half wave plate of quartz is given by

$$t = \frac{\lambda}{2(\mu_E-\mu_0)}$$

Here $\lambda = 5893 \times 10^{-8}$cm., $\mu_E = 1.553$ and $\lambda_0 = 1.544$

$$t = \frac{5890\times 10^{-8}}{2(1.533-1.544)}$$

$$= \frac{5890\times 10^{-8}}{2\times 0.009} = 3.27 \times 10^{-3} \text{ cm.}$$

(b) The thickness which would give the same result are t, 3t, 5t, ...

$$= 3.27 \times 10^{-3}\text{ cm.}, 3 \times 3.27 \times 10^{-3}\text{ cm.}, 5 \times 3.27 \times 10^{-3}\text{ cm}, ...$$

Ex. 13. The values of μ_E and μ_0 for quartz are 1.5508 and 1.5418 respectively. Calculate the phase retardation for $\lambda = 5000Å$ when the plate thickness is 0.032 mm.

Sol. The path difference between the ordinary and extraordinary rays for quartz plate of thickness t is given by

$$\text{path difference} = (\mu_E - \mu_0).\, t.$$

The phase retardation

$$= \frac{2\pi}{\lambda} \times \text{path difference}$$
$$= \frac{2\pi}{\lambda}(\mu_E - \mu_0)t.$$

Here $\mu_E = 1.5508, \mu_0 = 1.5418, t = 0.032\text{mm.} = 0.0032 \text{ cm. and } \lambda = 5000 \times 10^{-8} \text{ cm.}$

∴ Phase retardation

$$= \frac{2\pi}{5000 \times 10^{-8}} \times (1.5508 - 1.5418) \times 0.0032$$

$$= \frac{2\pi \times 0.009 \times 0.0032}{5000 \times 10^{-8}}$$
$$= 1.152\pi \text{ radians.}$$

Ex.14. A plate of thickness 0.020 mm is cut from calcite with optic axis parallel to the face. Given $\mu_0 = 1.648$ and $\mu_E = 1.481$ (ignoring variations with wavelength), find out those wavelengths in the range 4000 Å to 7800 Å for which the plate behaves as a half wave plate, and also those for which the plate behaves as a quarter wave plane.

Sol. The given calcite plate will behave as a half wave plate for wavelength λ if

$$(\mu_0 - \mu_E)t = (2n-1)\frac{\lambda}{2} \qquad \text{where } n = 1,2,3, \ldots$$

i.e.

$$\lambda = \frac{2\,(\mu_0 - \mu_E)t}{(2n-1)}$$

$$= \frac{2\,(1.648 - 1.481) \times 0.0020}{(2n-1)} \text{cm.}$$

$$= \frac{66800}{2n-1} \times 10^{-8} \text{cm.} = \frac{66800}{2n-1} Å$$

Substituting $n = 5,6,7,8$ we get $\lambda = 7422$ Å, 6073 Å, 5138 Å ,4453 Å in the visible region. These are the required wavelengths for which the given calcite plate behaves as a half wave plate.

The given calcite plate will behave as a quarter wave plate for wavelengths λ if

$$(\mu_0 - \mu_E)t = (2n-1)\frac{\lambda}{4} \qquad \text{where } n = 1,2,3, \ldots$$

i.e.

$$\lambda = \frac{4(\mu_E - \mu_0)t}{(2n-1)}$$
$$= \frac{4(1.648 - 1.481) \times 0.0020}{(2n-1)} \text{cm.}$$
$$= \frac{133600}{2n-1} \times 10^{-8} \text{cm} = \frac{133600}{2n-1} Å$$

Substituting n = 10,11,12,13,15,16,17 we get λ=7032 Å, 6362 Å, 5807 Å ,5344 Å, 4948 Å, 4607 Å,4310 Å,4048 Å in the visible region. These are the required wavelengths for which the given calcite plate behaves as a half wave plate.

Ex.15. For a given wavelength 1 mm of quartz cut perpendicular to optic axis rotates the plane of polarisation by 18°. Find for what thickness will no light of this wavelength be transmitted when the quartz plate is interposed between a pair of parallel nicols?

Sol. The polariser and analyser are parallel in this case, so the light is completely transmitted through the analyser. If no light is allowed to be transmitted through the analyser, the plane of vibration (or plane of polarisation) of light passing through quartz must be rotated through 90°.

It is given that 1 mm of quartz produces a rotation of 18°.

As the rotation is directly proportional to the thickness of the quartz, therefore, thickness of the quartz required to produce a rotation of 90° $= \frac{1}{18} \times 90 = 5\text{mm}$.

Based On Specific Rotation

Ex.16. A 20 g of cane sugar is dissolved in water to make 50 c.c. of solution. A 20 cm, length of the solution causes + 53°30′ optical rotation. Calculate the specific rotation (α).

Sol. The specific rotation is given by

$$\alpha = \frac{10\theta}{lc}$$

Here $\theta = 53°30' = 53.5°, c = \frac{20}{50} = 0.40\frac{g}{c.c}, l = 20$ cm.

$$\therefore \quad \alpha = \frac{10\times 53.5}{0.40\times 20} = 66.9°\,\text{deg}\,\text{dm}^{-1}(\text{g/c.c.})^{-1}.$$

Ex.17. A tube of sugar solution 20 cm. long is placed between crossed nicols and illuminated with light of wavelength 6×10^{-5}cm. If the optical rotation produced is 13° and the specific rotation is 65° $\text{decimetre}^{-1}(\text{g/cm}^3)^{-1}$, determine the strength of the solution.

Sol. The specific rotation (α) is given by

$$\alpha = \frac{10\theta}{lc}$$

Here $\alpha = 65°, \theta = 13°, l = 20$cm.

$$65° = \frac{10\times 13°}{20c}$$

$$\therefore \quad c = \frac{10\times 13°}{20\times 65°} = \frac{1}{10} = 0.1\ \text{g/cm}^3.$$

Ex. 18. On introducing a polarimeter tube of 25 cm. long containing a sugar solution of unknown strength it is found that the plane of polarisation is rotated through 10°. Find the strength of the solution in g/cm^3. Given specific rotation of sugar solution is 60° per decimetre per unit concentration.

Sol. The specific rotation (α) is given by

$$= \frac{10\theta}{lc} \text{ where } l \text{ is in cm.}$$

$$\therefore \quad c = \frac{10\theta}{l\alpha}$$

Here, $\theta = 10°, l = 25$ cm, $\alpha = 60°\,\text{deg}(\text{decimeter})^{-1}(\text{g/cm}^3)^{-1}$

$$\therefore \quad c = \frac{10\times 10}{25\times 60} = \frac{1}{15} = 0.0667\ \text{g/cm}^3.$$

*****Review Questions*****

Based On Polarisation

1. What do you mean by 'Polarisation'?
2. Distinguish between polarised and unpolarised light?
3. Is ordinary light polarised or unpolarised?
4. Define to terms 'plane of vibration' and 'plane of polarisation'.
5. Can sound waves be polarized? Explain.
6. Comment on the statement "Polarisation conclusively demonstrates that" light waves are transverse".
7. Mention various methods of producing plane polarised light and describe the one which consider to be the best.

Based On Brewster's law

8. What is Brewster's law?
9. What is the relation between refractive index and angle of polarisation?
10. State Brewster's law and use it to prove that when light is incident on a transparent substance at the polarising angle, the reflected and refracted rays are at right angles to each other.
11. Calculate the polarisation angle for crown glass ($\mu = 1.520$), and flint glassμ=1. 650). [Ans. 56°40', 58°47']

Based On Malus law

12. State Malus law?
13. Unpolarised light falls on a polarizing sheet. Show that the intensity of the transmitted plane polarized light is half the intensity of the incident unpolarised light.
14. Unpolarised light intensity I_o falls on two crossed polarising sheets so that no light is transmitted. If a third polarising sheet is placed between them, can light be transmitted. Explain
15. 2. Two polarizing sheets have their polarizing directions parallel so that the intensity of the transmitted light is maximum. Through what angle must be either sheet be turned if the intensity is to drop one-half? [Ans. ± 45°, ± 135°]
16. A polariser and an analyser are oriented so that the amount of light transmitted is maximum. To what percentage of its maximum value is the intensity of the transmitted light reduced when the analyser is rotated through (i) 30°, (ii) 45°, (iii) 60° and (iv)90°.[Ans. (i) 7.5%, (ii) 50%, (iii) 25% (iv)0]
17. 4.A polariser and an analyser are oriented so that the amount of light transmitted is maximum. How will you orient the analyser so that the transmitted light is reduced to (i) 0.5, (ii) 0.25, (iii) 0.75, (iv) 0.125, (v) 0 of its maximum value? [Ans. (i) 45°, (ii) 60°, (ii) 30°, (iv) 69°, (v) 90°]
18. Two Nicol prisms have their planes parallel to each other. One of the two Nicol prisms is then turned so that its principal plane makes an angle of 40° with the other. Calculate the percentage of the light originally transmitted by the second Nicol prism is now transmitted by it. [Ans. 59% intensity]
19. Two Nicols are oriented with their principal planes making an angle of 60°. What percentage of incident unpolarised light will pass through the system. [Ans. 12.50%]

Based On Double Refraction

20. What do you mean by "double refraction"?
21. (a)Explain the phenomenon of double refraction in calcite or quartz.
22. If a mark on a piece of paper is seen through a calcite crystal, two images are, in general, observed; however, if the crystal is oriented property, only one image can also be observed, Explain this phenomenon.
23. What do you mean by "ordinary ray" and "extraordinary ray"?

Based On Nicol Prism

24. What is a Nicol Prism?
25. What do you mean by the terms "Polariser" and 'Analyser'?
26. Name the crystal for which $\mu_E > \mu_o$ and for which $\mu_E < \mu_o$ where μ_E = refractive index of crystal for extraordinary (E)-ray, μ_o = refractive index of crystal for ordinary (o)-ray.
27. What do you understand by double refraction? Describe the construction, working and use of a Nicol prism.
28. Describe the construction of a Nicol's prism. Explain how it can be used as a polariser and as a analyser. Would a similar prism prepared from quartz serve a similar purpose?
29. Describe the construction of a Nicol Prism and show how the ordinary ray is quenched.

Based On Retardation Plates

30. Define a quarter wave plate and a half wave plate.
31. What is quarter wave plate? Describe its method of construction and use.
32. Deduce its thickness for a given wavelength in terms of refractive indices.
33. Explain how it is used with a Nicol prism for the analysis of polarised light?
34. Calculate the minimum thickness of a quarter wave plate of calcite for 5460 Å. The principal indices for calcite are 1.662 and 1.488.
35. Calculate the thickness of quarter wave plate for light of wavelength 5800 Å for which extra-ordinary and ordinary refractive indices are 1. 5708 and 1.5818 respectively. [Ans. 1.32×10^{-3} cm.]
36. Calculate the thickness of a quarter wave plate for light of wavelength 6000 Å (The refractive indices for ordinary and extraordinary rays are 1.544 and 1.533 respectively).[Ans. 1.66×10^{-3} cm.]
37. A quarter wave plate is to be designed from a crystal for wavelength 6×10^{-5} cm. If $\mu_0 - \mu_E$ = 0.0057, calculate the thickness of the plate. [Ans. 2.63×10^{-3} cm.]
38. 13.Find the thickness of a quarter wave plate when the wavelength of light is equal to 5890 A^0 and $\mu_0 = 1.55$, $\mu_E = 1.54$. [Ans. 2.25×10^{-3} cm.]
39. Calculate the thickness of (i) a quarter wave plate, (ii) a half wave plate. Given that, μ_E = 1.553, $\mu_0 = 1.544$ for $\lambda = 5000$ Å. [Ans. (i) 1.39×10^{-3} cm. (ii) 2.78×10^{-3} cm.]
40. The values of μ_E and μ_o for quartz are 1.5508 and 1.5418 respectively. Calculate the phase retardation for $\lambda = 5 \times 10^{-5}$ cm. when the plate thickness is 0.035 mm. [Ans. 1.152 π radians]
41. Calculate the thickness of a doubly refracting crystal required to introduce a path difference $\lambda/2$ between the ordinary and extraordinary rays when $\lambda = 6000$ Å, $\mu_0 = 1.550$ and $\mu_E = 1.540$. [Ans. 3×10^{-3} cm.]
42. A quarter wave plate is meant for λ = 5893 Å. Calculate the phase retardation which the plate will show for λ = 4358 Å. The variation of refractive indices with λ may be assumed to be negligible. [Ans. 0.67π]

43. Calculate the thickness of a quartz plate for the Fraunhoffer C - line (λ for C-line is 6563 Å) for which the extra ordinary and ordinary refractive indices of quartz 1.55085 and 1.54181 respectively. [Ans. 0.00363 cm]
44. For a given wavelength 1 mm of quartz cut perpendicular to the optic axis rotates the plane of polarization by 20^0.Find for what thickness will no light of this wavelength be transmitted when the quartz plate is interposed between parallel Nicols. [Ans. 4.5 mm]

Based on Production and Analysis of Light

45. What is meant by circularly and elliptically polarised light?
46. How would you produce and detect the following with the help of a Nicol prism and a quarter wave plate:
47. (i) Plane polarised, (ii) circularly polarised, (iii) elliptically polarised light.
48. What is polarised light? How will you produce and detect plane, elliptically and circularly polarised light?
49. Analyse mathematically the production of plane, circularly and elliptically polarised light by superposition of light moves.
50. What will be the state of polarisation of the emergent light,
(a)When linearly polarized light passes through a half wave plate whose optic axis is inclined to the plane of vibration of incident light.
(b)Linearly polarized light passes through a half wave plate whose optic axis is (i) parallel (ii) perpendicular to the plane of vibration of incident light.
(c)A beam of circularly polarised light is passed through a quarter wave plate?
(d)A beam of plane polarised light is passed through a quarter wave plate?
(e)A beam of elliptically polarised light is passed through a quarter wave plate whose optic axis is parallel to one of the two axes of the ellipse.
(f)A beam of plane polarised light is passed through a quarter wave plate such that the vibrations of the incident plane polarised light make an angle of 45° with the optic axis of the plate.[Ans. (a) plane polarised, (b) plane polarised, (c) plane polarised, (d) elliptically polarised, (e)plane polarised, (f) circularly polarised.]

Based On Optical Rotation, Specific Rotation

51. What is meant by optical activity?
52. What do you mean by dextro-rotatory and laevo-rotatory optically active substances? Give one example of each.
53. What is meant by rotatory dispersion'?
54. Define 'specific rotation'.
55. What is the effect of wavelength of light on 'angle of rotation'.
56. State the laws of rotatory polarisation. Give Fresnel's hypothesis for rotatory polarisation and derive a formula for the optical rotation produced by quartz. Give the experimental verification for this formula.
57. Describe and explain the phenomenon of optical rotation in quartz.
58. Discuss phenomenon of the rotation of the plane of polarisation of light by optically active materials. Give the necessary theory. Show that the rotation of plane of vibration is given by $\frac{\pi d}{\lambda}(\mu_A - \mu_C)$

59. where μ_A and μ_C are the refractive indices of the crystal in the direction of the optic axis for anti-clockwise and clockwise circularly polarised light and d is the thickness of the crystal plate.
60. Explain how polarised light is produced in polarimeter. Why is an arrangement of two crossed Nicols alone not preferable in a polarimeter? Describe the use of a polarimeter for determining the specific rotation of cane sugar solution.
61. What do you understand by rotatory polarisation and rotatory dispersion? Describe a half shade polarimeter and explain its limitations?
62. The specific rotation of quartz for 5893 Å is 21.7^0 per mm. What thickness of quartz cut perpendicular to its optic axis and inserted between parallel Nicols, will allow no light of this wavelength to be transmitted? [Ans. 4.147 mm]
63. The rotation in the plane of polarization (λ = 5893 Å) in a certain substance is 10^0 per cm. Calculate the difference between the refractive indices for right and left circularly polarized light in the substance. [Ans. 8.185×10^{-7}]
64. 80g of impure sugar when dissolved in a litre of water gives an optical rotation of 9.9^0 when placed in a tube of length 20 cm. If the specific rotation of sugar is 66°, find the percentage purity of the sugar sample. [Ans. 93·75%]
65. How will you construct a biquartz plate for finding the angle of rotation of wavelength 5890A^0. (Specific rotation of quartz for wavelength 5890 Å is 21.72 deg mm^{-1}, thickness of quartz = 4.14 mm)
66. A tube 20 cm long filled with an aqueous solution containing 15 g of cane-sugar per 100 c.c. of solution, is placed in the path of plane polarised light. Find the angle of rotation of the plane of polarisation. (Specific rotation of sugar is 66°) [Ans. 19.8^0]
67. A sugar solution in a tube of length 20 cm produces optical rotation of 13°. The solution is then diluted to one-third of its previous concentration. Find the optical rotation produced by 30 cm. long tube containing the diluted solution. [Ans. 6.5°]
68. A 15 cm. tube containing cane sugar solution (specific rotation 66°) show optical rotation 7°. Calculate the strength of the solution. [Ans. 0.07 g/cc]

****

Unit- IV

Theory of relativity: The Michelson-Morley Experiment and the speed of light; Absolute and Inertial frames of reference, Galilean transformations, the postulates of the special theory of relativity, Lorentz transformations, time dilation, length contraction, velocity addition, mass energy equivalence. Invariance of Maxwell's equations under Lorentz Transformation.

Chapter

6

Special Theory of Relativity

Introduction

In 1905, Albert Einstein published his pioneering work in the field of special theory of relativity and mass-energy equivalence. According to his work, space and time are not absolute rather relative. The relativeness of size, shape, weight, velocities and so many physical quantities are very simple to understand in context with Newtonian Mechanics, but when the measurements are done at a very high velocities (comparable to the velocity of light), the Newtonian Mechanics under Galileo transformation does not hold good. This is due to the fact that the relative motion of any object with respect to velocity of light could not be treated under Galileo transformations. This leads to "Einstein's Theory of Relativity". This area was being explored in earlier centuries too but Galileo and Newton were the first to correctly explain the version of classical relativity wherein space and time were considered as absolute and separate entities. Einstein put an effort to summarize the fact that classical laws are inapplicable for the bodies moving closely with the velocity of light. The analysis of length, time and mass at high velocities yielded different results. The concept of modern relativity by Albert Einstein thus verified and developed two important classes of relativity:" Special relativity" deals with observers moving at constant velocity and "General relativity" deals with observers undergoing acceleration. Both the theories are well explained considering the frame of references. First, we will explain the frame of references and then we shall deal here only with special theory of relativity.

6.1 Experimental Background of Theory of Relativity (Michelson-Morley Experiment)

Objective of the Experiment

Initially, it was thought that a transparent medium is filled everywhere in free space which produces no resistance for any type of propagation. It is called ether and is considered as best absolute frame of reference. Michelson-Morley experiment was an attempt to confirm the existence of stationary ether or to locate an absolute frame of reference for space and time.

Principle

Michelson-Morley experiment is performed using Michelson interferometer which is based on the interference due to division of amplitude. Actually, the principle of the experiment lies in noting the fringe shift due to the difference in time taken by the light in travelling along and opposite to the direction of motion of the earth.

Construction

The main apparatus used in experimental arrangement is Michelson interferometer, shown in Fig.6.1. It consists of an extended source of light, S, a half-silvered plate P which is inclined at an angle of 45° to the beam of light, two mirrors M_1 and M_2 highly silvered on their front surfaces so as to avoid multiple internal reflections and nearly hundred percent reflection.

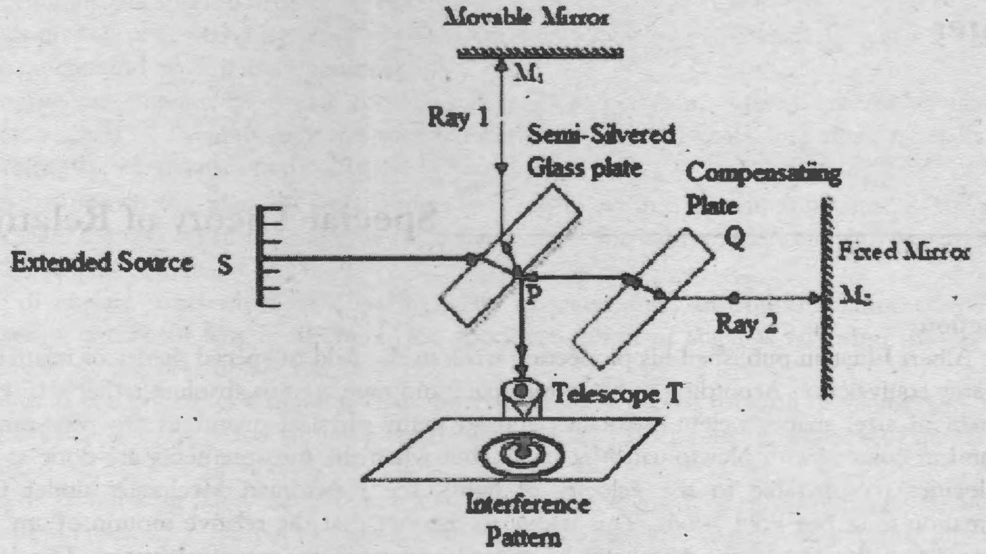


Fig. 6.1

When beam of light falls from a source at plate P, the semi-silvered plate divides light into reflected and transmitted beams of nearly equal amplitudes. The reflected ray '1' moves towards M_1 and reflected back to P and the transmitted beam '2' moves towards M_2 and is reflected back from mirror M_2 to plate P. The two beams '1' and '2' can then interfere constructively or destructively depending on their phase difference and a fringe pattern can be observed in the telescope 'T'.

Experiment

Let us first consider that the medium ether around the apparatus is at rest.

Let, c = the velocity of light

L = the distance PM_1 or PM_2

t_1 = time taken by the light to travel distance PM_1 and then M_1P

t_2 = time taken by light to travel distance PM_2 and then M_2P.

The whole apparatus as situated on the Earth is moving with Earth's velocity in the direction of propagation of initial beam of light. Due to motion of the earth, the reflections at the two mirrors M_1 and M_2 do not take place at M_1 and M_2 but at M_1' and M_2' respectively. Let us calculate the time difference of two light beams for two different paths during the motion of the whole system.

Now, reconsidering Fig. 6.1, as Fig. 6.2, when we take the motion of the earth into consideration and calculate time T_1 and T_2 i.e. time travelled by the ray towards mirror M_1 and back to the glass plate and towards mirror M_2 and back to the glass plate respectively.

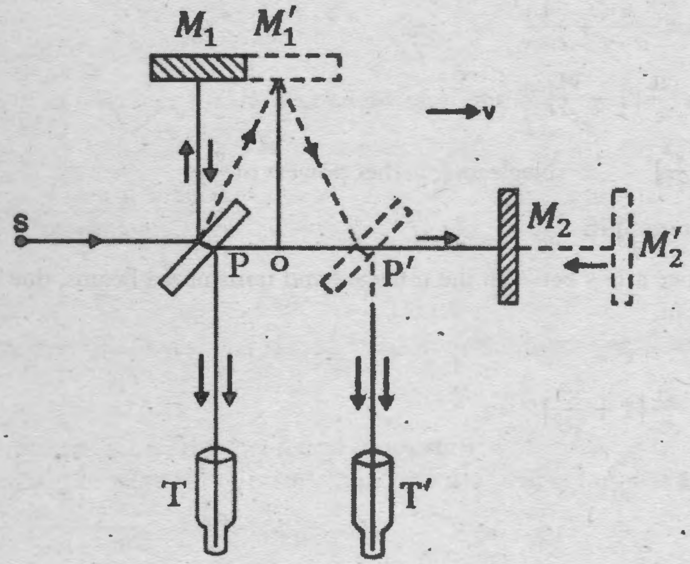


Fig. 6.2

Calculation Of Time T_1

From $\Delta\, PM_1'P'$

$$(PM_1')^2 = (PO)^2 + (M_1'O)^2$$

From the geometry of the Fig. 6.2, we have;

$PM_1' = c \times t = ct$ and $PO = v \times t$, (v being the velocity of apparatus)

As, $M_1'O = L$, then;

$$(ct)^2 = (vt)^2 + L^2$$

$$t^2(c^2 - v^2) = L^2$$

Or $$t = \frac{L}{\sqrt{c^2 - v^2}}$$

Time taken to cover path PM_1P' is $2PM_1' = 2t$ (as $M_1'P' = PM_1'$)

So, the total time taken;

$$T_1 = \frac{2L}{\sqrt{c^2 - v^2}} = \frac{2L}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2L}{c}\left(1 - \frac{v^2}{c^2}\right)^{-1/2}$$

$$= \frac{2L}{c}\left[1 + \frac{v^2}{2c^2}\right]$$ (neglecting higher powers of v/c)

Calculation of T_2

Time taken by the ray '2', to travel a distance L towards $PM_2' = \frac{L}{c - v}$

And the time taken by the ray to travel a distance L along $M_2'P = \frac{L}{c+v}$

Thus, total time, $T_2 = \frac{L}{c-v} + \frac{L}{c+v}$

$$= \frac{2Lc}{c^2 - v^2} = \frac{2L}{c}[1 - \frac{v^2}{c^2}]^{-1}$$

$$= \frac{2L}{c}[1 + \frac{v^2}{c^2}] \qquad \text{(Neglecting higher powers of } \frac{v}{c}\text{)}$$

Calculation Of Fringe Shift

Let the time difference arises between the reflected and transmitted beams, due to the motion of the Earth system is Δt. So;

$$\Delta t = T_2 - T_1$$

$$= \frac{2L}{c}\left[1 + \frac{v^2}{c^2}\right] - \frac{2L}{c}[1 + \frac{v^2}{2c^2}]$$

$$= \frac{2L}{c}[1 + \frac{v^2}{c^2} - 1 - \frac{v^2}{c^2}]$$

$$= \frac{2L}{c}[\frac{v^2}{c^2} - \frac{v^2}{2c^2}]$$

$$\Delta t = \frac{2L}{c} \frac{v^2}{2c^2} = \frac{Lv^2}{c^3}$$

The equivalent path difference between the two light beams is given as;

$$\Delta d = c\,\Delta t = c \times \frac{Lv^2}{c^3} = \frac{Lv^2}{c^2}$$

It means if the physical distance of the two mirrors is exactly same and the relative velocity between earth and ether is taken into consideration then there should be a shift in the interference fringe pattern by an amount of $\frac{Lv^2}{c^2\lambda}$, calculated as below;

$$\Delta d = \Delta N.\lambda \qquad \text{(Here } \Delta N \text{ is the numbers of the shifted fringes)}$$

$$\text{Or } \Delta N = \frac{\Delta d}{\lambda} = \frac{\textbf{change in path difference}}{\textbf{wavelength}}$$

$$\Delta N = \frac{Lv^2}{c^2\lambda}$$

But it is difficult to measure this shift as earth can never be at rest. Thus, to measure this, the whole apparatus was turned to 90° and repeat the experiment. This interchanged the roles of the two beams i.e. the path that originally required the time t_1 for the light to pass through, now requires the time t_2 and vice a versa. The new time difference will be;

$$\Delta t' = \frac{2L}{c}\left[1 + \frac{v^2}{2c^2}\right] - \frac{2L}{c}[1 + \frac{v^2}{c^2}]$$

$$= \frac{2L}{c}\left[-\frac{v^2}{2c^2}\right] = \frac{-Lv^2}{c^3}$$

And the corresponding path difference is;

$$\Delta d' = c\Delta t' = \frac{-Lv^2}{c^2}$$

Thus, the path difference between two beams, for the two different positions of the apparatus is,

$$\Delta D = \Delta d - \Delta d' = \frac{Lv^2}{c^2} - \left(\frac{-Lv^2}{c^2}\right)$$

Or $\Delta D = \frac{2Lv^2}{c^2}$

So, taking one position as reference for the second position, the total number of fringes shifted will be;

$$\Delta d = \Delta N.\lambda$$

Or $\Delta N = \frac{\Delta d}{\lambda} = \frac{\text{change in path difference}}{\text{wavelength}}$

$$\Delta N = \frac{2Lv^2}{c^2\lambda}$$

ΔN, are the total numbers of fringes, that should be seen in the telescope, for the two positions of the apparatus.

For a particular experimental setup, the theoretical fringe shift was calculated by taking $\lambda = 5000\text{Å}$, $L = 10$ m, $v = 3 \times 10^4$m/s (orbital speed of earth around the sun) that comes out to be;

$$\Delta N = \frac{2 \times (10\text{m}) \times (3 \times 10^4 \text{m/s})^2}{(3 \times 10^8 \text{m/s})^2 \times (5 \times 10^{-7}\text{m})} = 0.4 \text{ fringes}$$

But in the telescope no fringe shift was observed.
Since, the number of fringes shifted is proportional to L so to get more accurate results, the distance L was increased by Michelson and Morley by multiple reflections through a system of mirrors but again no shift of fringes was observed in the telescope. The result of this experiment was termed as the negative result as theoretical result never observed practically.

6.1.1 Negative Result and Its Explanation

The theoretically calculated value of fringe shift indicates that if ether is present everywhere on the earth and there is a relative motion between earth and ether, then there must be a fringe shift of 0.4 fringes in the telescope of the interferometer. It was also be thought that at the time of experiment the earth was at rest relative to ether due to the motion of solar system as a whole so the experiment was repeated after six months again when motion of the earth was considered again but the fringe shift could not be detected. Some other physicist tried the experiment day and night (as earth spins about its axis) during all seasons of the year (as the earth rotates about the sun) but could not find any fringe shift. In spite of all the efforts and arrangements to obtain the theoretical value of fringe shift, interferometer did not show any fringe shift. This contradiction between theoretical and experimental findings is called **negative results** of Michelson-Morley experiment.

The negative results of Michelson - Morley experiment was such a blow to ether hypothesis that the experiment was repeated by many scientists over a 50-year period and was one of the greatest puzzles of physics in the 19th century. Later on, it was thought that the negative result of the experiment ($\Delta N = 0$) can be explained if;

(1) Galilean transformation is not taken into account or in other words something should be modified in this transformation.
(2) Velocity of light should be assumed to be same in all inertial frames whether the light was moving with the earth or against it.

6.2 Speed of Light: Absolute Frame of Reference

Though, Newton's formulation of relativity did explain the existence of absolute space but classical mechanics was unable to provide any criterion for detecting an absolute reference frame. The answer to this came later with the formulation of theory given by Maxwell. It was believed that Maxwell's equation must obey Newtonian relativity i.e., Maxwell's equations must be the same in all inertial frames of reference. Unfortunately, it was not true and was found to be completely in different forms in different frames of references e.g., we can think of a simple form of Maxwell's equations of electromagnetic waves, in one frame of reference (say S) but it seems to be more complicated in another reference frame (say S') which is moving with a constant velocity with respect to first.
We have equation for electromagnetic waves in the form of the general wave equation as:

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = 0$$ (Where 'c' is the speed of electromagnetic waves in a frame at rest).

For a frame S', which is moving with a constant velocity v_x along x direction, the equation under Galilean transformation becomes;

$$\frac{\partial^2 E'}{\partial x'^2} - \frac{1}{c^2}\frac{\partial^2 E'}{\partial x'^2} - \frac{2\,v_x}{c^2}\frac{\partial^2 E'}{\partial x'\,\partial x'} - \frac{v_x}{c^2}\frac{\partial}{\partial x'}\left[v_x\frac{\partial E'}{\partial x'}\right] = 0$$

And seems to be not invariant under Galilean transformation.
One of the important consequences of the electromagnetic theory as formulated by James Clerk Maxwell is the consideration of light as an electromagnetic wave travelling in a vacuum with a constant speed given as:

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 2.997928 \times 10^8 \text{ m/sec}$$

This implies the speed of light must have a constant value 'c' with respect to all inertial frames, independent of the motion of the light source but in case of Newtonian or classical relativity, a particle has less velocity in a system which is moving in the direction of velocity of particle than a system at rest in case of Galilean transformation. This implies the two frames of reference one at rest and the other moving, experience the difference of velocity of light in two different systems. Maxwell's theory and Newtonian principle thus contradicts each other.

Thus, if we have to accept the electromagnetic theory, we need to modify our concept of space and time. There should be a frame of reference with respect to which the speed of light must be constant, or the transformation equations must be modified so as to be different from that given by Galilean.

6.3 Frame of Reference

The event can be defined as something which happens at an instant of time at a particular point in space or to be more precise which happens over a localized time interval and localized region in space. It can also be considered as another name for a point in space and time wherein this point can be determined by knowing the spatial co-ordinates (x, y, z) of the point in space and reading of a clock at that point (say at time t).
Thus, for different reference frames, an event can have different co-ordinates. One can even relate the co-ordinates of events in one reference frame to the co-ordinates of this same event in another reference frame. This can be explained by using a certain set of equations known as Galilean

transformation equations in Newtonian Physics and Lorentz transformation equations in special theory of relativity.

6.4 Inertial and Non-Inertial Frames of Reference

6.4.1 Inertial Frame of Reference

A reference frame in which Newton's law of inertia is valid is called an inertial frame of reference – this means a body at rest will remain at rest and the body in motion will remain in motion unless an external force acts on it. It may also be treated as a non-accelerated frame of reference.
The motion of a body though makes no sense unless it is determined with respect to some well-defined co-ordinate system or a frame of reference with respect to which the velocity of a body can be measured i.e. a proper choice of a co-ordinate system is always needed in order to define the motion of body. In case of an inertial reference frame, the acceleration is zero such that it moves with a constant velocity. Any reference frame moving with a constant velocity relative to another inertial frame is also an inertial frame of reference.

Consider any co-ordinate system relative to which a body in motion has co-ordinates (x, y, z). The co-ordinates of the object relative to the assumed co-ordinate system are functions of time such that the mathematical form for Newton's first law can be written as:

$$\frac{d^2x}{dt^2} = 0,\ \frac{d^2y}{dt^2} = 0,\ \frac{d^2z}{dt^2} = 0$$

$$\frac{d^2x}{dt^2} = 0,\ \frac{d^2y}{dt^2} = 0,\ \frac{d^2z}{dt^2} = 0 \qquad \text{(Since, the body is not subjected under external force)}$$

$$\frac{dx}{dt} = u_x,\ \frac{dy}{dt} = u_y,\ \frac{dz}{dt} = u_z$$

Where, u_x, u_y, u_z are the components of velocity in x, y, z direction respectively.

When there is no external force applied, the velocity components are constant and body continues moving with a uniform velocity.

The proper choice of a frame of reference or co-ordination system with respect to which the body is at rest or in uniform motion is thus essential. E.g., reference frames fixed on the earth can be considered as inertial frames.

6.4.2 Non-Inertial Frame of Reference

The reference frames where the law of inertia does not hold true are called non-inertial reference frames. Newton's first law is not obeyed as the body appears to be accelerating without any net force acting on it. E.g., the terrestrial frames can be regarded as non- inertial frames, another example is a rotating frame or carousel.

6.5 Galilean Transformation for Space and Time

In order to compare mathematically the laws of physics in different inertial reference frames we need to relate the co-ordinates of an event as observed in one inertial frame to the co-ordinates of the same event as observed by an observer in the other. For this, we need to derive some transformation equations mathematically showing consistency of Newton's law with that to the principle of relativity. It is but obvious that the two given frames have constant relative velocity so as to make law of inertia valid. We can here try to find out now position, velocity and acceleration in one frame corresponding to the other.

In order to derive these transformation equations, consider S and S' to be two different inertial frames of reference with S' moving with velocity v along x direction. Let the two frames of reference be parallel to each other such that $\mathbf{x'}$ is parallel to x axis, $\mathbf{y'}$ parallel to y and $\mathbf{z'}$ parallel to z axis.

Case I

When the second frame (here $\mathbf{S'}$) is moving relative to first frame (S) along positive x axis with constant velocity v.

Let us consider S' frame is moving relative to S (at rest) with speed v along x axis in Fig.6.3. The origins of two frames of reference are so chosen such that they coincide at $t = \mathbf{t'} = 0$. This defines the absolute time concept and an implicit assumption in Newtonian Mechanics wherein time flows uniformly without reference to anything external. Thus, in all the inertial systems, clocks will measure same time at any instant and hence the same time interval between two events.

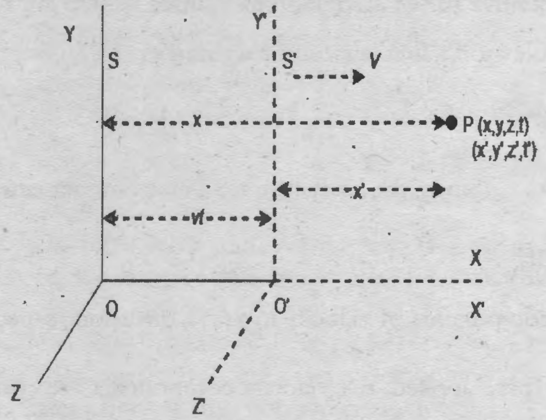


Fig. 6.3

If suppose an event E takes place at any particular time and the observations are taken by the observers in two different reference frames, the two can be related as follows:

$$\left.\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned}\right\} \qquad (1)$$

The set of equations (1) together are known as Galilean transformation equations and signifies the relation between the co-ordinates of an event in one inertial frames S with those of the co-ordinates of same event as measured in another frame S' moving with a constant velocity.

6.5.1 Inverse Galilean Transformation Equations

(1) These set of equations can be obtained by taking primed variable in terms of unprimed variables.

(2) By considering a case wherein S will be moving with a velocity –v with respect to S′. The inverse transformation equations (2) below thus can be obtained by exchanging the primed and unprimed variables and by replacing v by –v.

By either of the ways we get:

$$\left.\begin{aligned} x &= x + vt \\ y &= y' \\ z &= z' \\ t &= t' \end{aligned}\right\} \qquad (2)$$

Case II

When the second frame S′, is moving along a straight line relative to first in any direction.

Again, let us consider two frames of reference where S′ is moving relative to S, in Fig.6.4 with a velocity $\vec{v}$:

$$\vec{v} = i\overrightarrow{v_x} + j\overrightarrow{v_y} + k\overrightarrow{v_z}$$

Where v_x, y_y and v_z are the velocity components along x, y and z axes.

Now, let O and O′ be the two different observers situated at origin S and S′ respectively observing the same event at point P. The co-ordinates of P relative to S is (x, y, z, t) and to that to S′ is (x′, y′, z′, t′) respectively. At, t = t′= 0, the origins of the two frames coincide with each other. Let after a time t, the frame S′is separated by frame S by a distance $v_x t$, $v_y t$ and $v_z t$ along x, y and z axes respectively as shown in the Fig. 6.4 below.

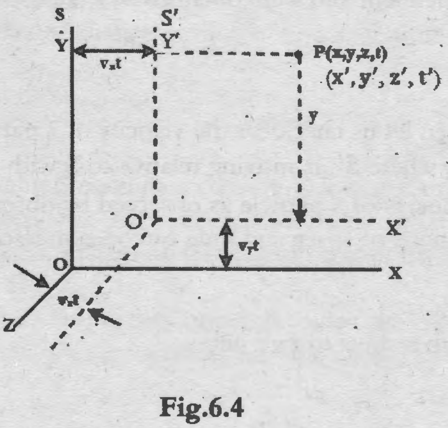


Fig.6.4

The two frames of reference can be related to each other as:

$$\left.\begin{aligned} x' &= x - v_x t \\ y' &= y - v_y t \\ z' &= z - v_z t \\ t' &= t \end{aligned}\right\} \qquad (3)$$

Case-III Vector Form of Galilean Transformation

Again, consider two reference frames S and S' where S' is moving with velocity $\vec{v}$ relative to S in Fig.6.5

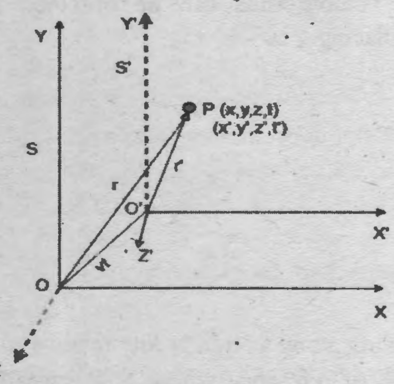


Fig. 6.5.

Initially the origins of two frames of reference coincide at $t = t' = \theta$. Let $\vec{r}$ be the position vector of an event at point P relative to origin O in system S and $\vec{r}'$ be the position vector of an event at point P relative to origin O' in system S' after time t such that $OO' = vt$.

Using triangle law of vector addition in $\Delta OO'P$:

$$\overrightarrow{OP} = \overrightarrow{O'P} + \overrightarrow{OO'}$$

$$\left.\begin{aligned} \vec{r} &= \vec{r}' + vt \\ \vec{r}' &= \vec{r} - vt \\ t' &= t \end{aligned}\right\} \qquad (4)$$

The above equations are Galilean transformation for space and time in vector form. The equations (1), (3) and (4) are all time dependent and were obtained by Galileo.

6.6 Galilean Invariance

6.6.1 Velocity of a Particle

If the positions transform, then let us check for the velocity of a particle. For this again consider two frames of reference S and S', where S', is moving relative to S with velocity, $\vec{v} = v_x i + v_y j + v_z k$. If r(t) and r' (t') are the co-ordinates of a particle as observed by observer in system S and S', then the Galilean transformation equations of space and time can be expressed as:

$r' = r - vt$ (5)(i)

$t' = t$ (5)(ii)

Differentiate equation5 (i) with respect to t we get

$$\frac{dr'}{dt} = \frac{dr}{dt} - v$$

$$\Rightarrow \quad u' = u - v \qquad (6)$$

Similarly, we can have;

$$u = u' + v$$

This is actually the addition of velocities under Galilean transformations.

6.6.2 Acceleration of a Particle

Let us now see the transformation for acceleration. Galilean transformation for velocity of particle is
$u' = u - v$

Differentiate equation (6) with respect to t, we get

$$\frac{du'}{dt} = \frac{du}{dt}$$

or $\frac{du'}{dt'} = \frac{du}{dt} \Rightarrow a' = a$ (as $dt' = dt$)

Where $\frac{du'}{dt'} = a'$, is acceleration of particle in S'

$\frac{du}{dt} = a$, is acceleration of particle in S

The above equation signifies that the acceleration in both the frames is same or in other words the acceleration is invariant under Galilean transformation. This is because the acceleration is the rate of change of velocity and velocities of the same particle measured in the two frames differ by a constant factor – the relative velocity of the two frames.

6.6.3 Newton's Law

If m be the mass of particle, then mass is also invariant under Newtonian mechanics, so

$F = ma$ (by Newton's II law of motion)

Since mass is invariant and acceleration too so force also remains invariant.

i.e., $F' = F$

$ma' = ma$

Thus, Newton's II law is valid in every inertial system and is invariant under Galilean transformation.

6.6.4 Conservation of Momentum

If $m\vec{a}$ is invariant, so will the rate of change of momentum as $\vec{p} = m\vec{v}$.

$$\frac{d\vec{p}}{dt} = m\frac{d\vec{v}}{dt} = m\vec{a}$$

This implies, in a collision, if total momentum is conserved in one frame (the sum of individual rates of change of momentum is zero) the same is true in all the inertial frames. Thus, the law of conservation of momentum is also invariant under Galilean transformation.

The above discussion leads to an important principle called as Newtonian or classical relativity principle.The principle explains the invariance of the laws of mechanics under Galilean transformation.

Newtonian relativity assumes space and time as absolute quantities. Their measurement does not vary from one inertial frame to another. The mass and force remain invariant but the position and velocity of an object are different in different inertial frames. This was accepted by all the physicists everywhere atleast until when Maxwell put together his famous set of equations.

6.7 Drawbacks of Newtonian Relativity/Galilean Transformation

- It was unable to determine whether or not a frame of reference is in a state of motion by any experiment using Newton's Laws. There is non-existence of any single preferred frame of reference. Observers in different inertial frames would discover the same mechanical laws. E.g., the results of an experiment conducted at a railway platform for a uniformly moving train will be same. An observer is unable to detect his motion in an inertial system by any mechanical experiment performed within the system.
- Nowhere do the Newton's laws depend on the velocity of a frame of reference relative to anything else even though Newton had given the postulate of "absolute space" i.e., a frame of reference which defines the state of "absolute rest" and with respect to which the motion "for anything" can be measured.

6.8 Variance of Velocity of Light - Need for Other Transformations Equations

To show whether the laws of electromagnetism are invariant under Galilean-Transformations, let us think an electromagnetic wave that propagates with velocity c in ether (initially vacuum was considered as medium called ether), with respect to stationary frame at rest (say S) in Fig.6.6.

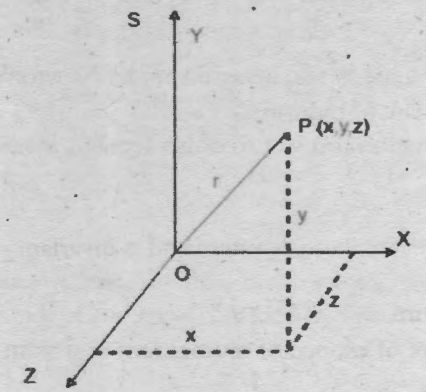


Fig. 6.6

The speed of electromagnetic wave is given by, $c = \frac{r}{t}$, where r is the distance travelled from the source of the wave to point P.

$$r = ct \tag{7}$$

or $$r^2 = c^2t^2 \tag{8}$$

$$r^2 - c^2t^2 = 0 \tag{9}$$

The radius r of the wavefront (spherical wavefront) can be given as:

$$r^2 = x^2 + y^2 + z^2 \tag{10}$$

Substituting equation (10) in (9) we get

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \tag{11}$$

Equation (11) shows the light wave as observed in S frame of reference.

For an observer which is moving at a speed v in a moving frame of reference S', the co-ordinates x' and t' can be related to x and t co-ordinates by Galilean transformation as:

$$\left.\begin{aligned} x &= x' + vt \\ y &= y' \\ z &= z' \\ t &= t' \end{aligned}\right] \qquad (12)$$

Substituting (12) in (11), we get:

$$(x' + vt) + y'^2 + z'^2 - c^2t'^2 = 0 \qquad (13)$$

or $x'^2 + 2x'vt + v^2t^2 + y'^2 + z'^2 - c^2t'^2 = 0$

or $(x' + vt) + y'^2 + z'^2 - c^2t'^2 = 0$

or $x'^2 + y'^2 + z'^2 - c^2t'^2 = -2x'vt - v^2t^2 \qquad (14)$

This implies that the equations (11) and (14) show variance of the velocity of light as observed in S and S' frame. Einstein then realized that something might be wrong with Galilean transformation that needs to be modified and proposed his well-known postulates for special relativity.

6.9 Einstein's Postulates of Special Theory of Relativity

In 1905, Albert Einstein developed Special Theory of Relativity. He dropped the concept of ether and assumption of an absolute frame of reference at rest. He dropped the idea of absolute motion and revised the classical theory in space and time. Einstein's thoughts and his pioneer work were explained in two important postulates;

Postulate 1. The laws of physics have the same form in all the inertial frames of reference.

Postulate 2. The speed of light in free space has a constant value c regardless of the relative velocity between source and observer.

Explanation of Postulate 1

It is the extension of the conclusion which is drawn from Newtonian mechanics, since velocity is not absolute but relative which can be explained from the failure of the experiments to determine the velocity of earth relative to ether. At constant velocity, the difference between rest and motion cannot be detected e.g., consider two brothers travelling in space in smoothly and uniformly moving rockets. If they are too far apart, they cannot say whether they are at rest or in motion. When the two come near each other they can be easily compared.

Explanation of Postulate 2

According to this postulate, the speed of light is same in all the directions, irrespective of whether the source of light is moving or stationary or whether the velocity of light is measured relatively to medium in which it travels or relatively to a moving observer. It is the greatest velocity that can be achieved and is the ultimate speed. The second postulate however explains the result of Michelson

Morley experiment. Thus, the theory based on the above two postulates applies to all the inertial frames and called as special theory of relativity.

6.10 Lorentz Transformations Equations for Space and Time

Einstein's second postulate is extremely important to distinguish the classical theory of relativity and Einstein's theory of relativity. The constancy of velocity of light in any frame of reference, thus demands a new set of transformation equations that can relate the positions and times of an object in two relatively moving frame of references. The following requirements thus should be fulfilled by these new set of transformation equations,

(1) The speed of light 'c' must be constant in any inertial frame.
(2) The new equations must not be based on "absolute time" and "absolute space".
(3) The transformation must be linear and for low velocities, i.e., for $v << c$, must reduce to Galilean transformation.

H.A Lorentz could satisfy the above requirements and introduced new set of transformation equations called "Lorentz Transformation Equations". Now let us derive these transformation equations;

Consider two inertial frames of reference S and S′ with S at rest and S′ moving with respect to S with a constant velocity v parallel along x axis shown in Fig.6.7. Let the observer at O and O′ observe any event P from S and S′ respectively.

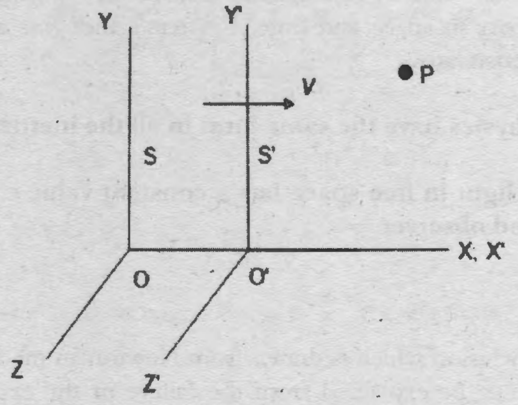


Fig. 6.7.

At $t = t' = 0$, the origins of the two frames coincide and an event P is a light signal say, having co-ordinates (x, y, z, t) and (x′, y′ z′, t′) as observed by observers O and O′ in S and S′ frames respectively.

At $t = 0$, the light pulse spreads out as a growing sphere and radius of wavelength will grow with speed c. For an observer in rest frame S, with co-ordinates (x, y, z, t) the equation of spherical surface whose radius grows at speed c is:

$$x^2 + y^2 + z^2 = c^2t^2 \quad \text{or} \quad x^2 + y^2 + z^2 - c^2t^2 = 0 \qquad (15)$$

Similarly, for an observer in S′ frame with co-ordinates (x′, y′ z′, t′), the equation of spherical surface is:

$$x'^2 + y'^2 + z'^2 = c^2t'^2 \quad \text{or} \quad x'^2 + y'^2 + z'^2 - c^2t'^2 = 0 \qquad (16)$$

('c' is kept constant in equation (15) and (16) i.e.; for two different inertial frames because of second postulate)

From (15) and (16):

$$x^2 + y^2 + z^2 - c^2t^2 = \lambda\left(x'^2 + y'^2 + z'^2 - c^2t'^2\right) \qquad (17)$$

Where λ, is any undetermined constant.

Here we have taken, $y = y'$ and $z = z'$, for the motion in one dimension, i.e., along x-axis, the value of λ comes out to be 1.

So, from (17) we get; $\quad x^2 - c^2t^2 = \left(x'^2 - c^2t'^2\right) \qquad (18)$

Consider, a suitable transformation equation that can relate x and x' as;

$$x' = \gamma(x - vt) \qquad (19)$$

Here, γ is independent of x and t.

The reasons for using the above relation are:

(1) The Galilean transformations are correct when the motion of the body is at low speed and the new transformation equation can reduce to Galilean equations.

(2) The transformation equation should be simple and linear.

Equation (19) thus, satisfies the above both the requirements. This states that if the position 'x' and velocity 'v' of a body are measured in the rest frame, then its position 'x'' in the moving frame can be determined using the first postulate of relativity, this equation must have the same form in the frame of reference at rest.

Thus, $\quad x = \gamma(x' + vt') \qquad (20)$

Where x' is the position of body in the moving frame at the time t'. The sign of v is changed to positive because the frame '2' moving to the right with a velocity 'v' as observed from a frame '1' at rest is equivalent to frame '2' at rest with frame '1' moving to the left with a velocity (-v). This equation shows the position x' and time t' of a body are measured in a moving frame, then its position x in the stationary frame is determined by equation (20).

With $\quad y = y'$ and $\quad z = z'$

Putting values of x' from (19) in (20), we have;

$$x = \gamma\,[\,\gamma(x - vt) + vt'\,]$$

$$\text{Or } t' = \gamma\left[t - \frac{x}{v}\left(1 - \frac{1}{\gamma^2}\right)\right] \qquad (21)$$

Substitute the value of x', from equation (19) and t', from equation (21) in equation (18), we get:

$$x^2 - c^2t^2 = \left[\gamma^2(x - vt)^2 - c^2\gamma^2\left\{t - \frac{x}{v}\left(1 - \frac{1}{\gamma^2}\right)\right\}^2\right]$$

$$\text{Or } x^2 - c^2t^2 - \gamma^2(x^2 - 2vxt + v^2t^2) + c^2\gamma^2\left\{t - \frac{x}{v}\left(1 - \frac{1}{\gamma^2}\right)\right\}^2 = 0 \qquad (22)$$

Equation (22) is an identity, so the coefficient of various powers of x and t must therefore vanish separately. Now equating the coefficients to x^2 to zero, we get;

$$1-\gamma^2+\frac{c^2\gamma^2}{v^2}\left(1-\frac{1}{\gamma^2}\right)^2=0$$

$$1-\gamma^2+\frac{c^2\gamma^2}{v^2}\left(1-\frac{2}{\gamma^2}+\frac{1}{\gamma^4}\right)=0$$

$$1-\gamma^2+\frac{c^2}{v^2}\left(\gamma^2-2+\frac{1}{\gamma^2}\right)=0 \tag{23}$$

Again, equating the coefficients of 'xt' to zero, we get;

$$2\gamma^2 v+c^2\gamma^2\left\{-\frac{2}{v}\left(1-\frac{1}{\gamma^2}\right)\right\}=0 \tag{24}$$

Or $\quad 2\gamma^2 v-\frac{2c^2\gamma^2}{v}\left(1-\frac{1}{\gamma^2}\right)=0$

Or $\quad (v^2-c^2)\,\gamma^2+c^2=0 \qquad (25)$

Equation (24) gives, $\quad \gamma=\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \qquad (26)$

From equation (19), $\quad x'=\gamma(x-vt) \qquad (27)$

From equation (21), $\quad t'=\gamma\left[t-\frac{x}{v}\left(1-\frac{1}{\gamma^2}\right)\right]$

$=\gamma\left[t-\frac{x}{v}\left(1-\frac{c^2-v^2}{c^2}\right)\right]$ (from equation 26)

$=\gamma\left[t-\frac{v^2x}{vc^2}\right]$

$$t'=\gamma\left[t-\frac{vx}{c^2}\right] \tag{28}$$

Substitute value of γ from equation (26) in equation (27) and (28), we have

$$x'=\frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}} \quad \text{and} \quad t'=\frac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \tag{29}$$

$$\left.\begin{aligned} x'&=\frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}}\\ y'&=y\\ z'&=z\\ t'&=\frac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \end{aligned}\right\} \tag{30}$$

These are Lorentz transformation equations.

If we exchange our frames of reference i.e., if we observe space – time co-ordinates of the event in frame S′ rather than in frame S, change relative velocity v to – v i.e., the frame S moves relative to system S′ with velocity v along negative direction of x axis where frame S′ moves relative to system S along positive direction of X axis.

$$\left.\begin{aligned} x &= \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}} \\ y &= y' \\ z &= z' \\ t &= \frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned}\right] \qquad (31)$$

Equations (31) are Inverse Lorentz Transformation equations.

Special Case:

If $v << c$, $\frac{v}{c} \ll 1$, Equation (30) reduces to

$$\left.\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z. \\ t' &= t\left(\frac{vx}{c^2} \ll t\right) \end{aligned}\right\} \qquad (32)$$

These are Galilean transformations. Thus, for low values of velocity v, the Lorentz transformations reduces to Galilean transformations. The Lorentz transformation holds true for all velocities while Galilean for only low velocities.

Following is the interpretation that can be drawn from the above transformation equations:

Equations (30) and (31) show the dependence of x' on x and similarly t' on t and vice a versa too. This indicates the dependence of both space co-ordinates and time of an event in one frame to the other frame. These have no absolute meaning independent of the frame of reference. Thus, space and time are very well connected.

- If $v >> c$, $\sqrt{1 - \frac{v^2}{c^2}}$ will be imaginary and so be the space and time co-ordinates which is physically unacceptable. Thus, in vacuum speed of an object cannot exceed the speed of light.
- If $v << c$, the Lorentz transformation equations reduces to the Galilean transformation equations as $\frac{v^2}{c^2} \approx 0$. The relativistic effects are observable only at high speeds.

6.11 Outcome of Lorentz Transformation Equations

Lorentz transformations are the consequences for the preservation of Einstein's postulates. These transformation equations also incorporate Galilean transformation equations as well as Newtonian Mechanics at low velocities. These are some of the biggest merits of these equations. There are some interesting but unusual (from Newtonian Mechanics point of view) findings of Lorentz Transformation Equations.

6.11.1 Length Contraction

The most interesting aspect of Lorentz transformation is that length has no meaning of the word "absolute". The length of an object depends on the motion relative to the frame of reference in which the measurement of length takes place.

Let us consider two reference frames S and S′. S is a stationary frame and S′ is moving relative with S at a constant velocity v along x axis. A rod is lying parallel along x axis in system S′at rest. Since the rod is stationary in S′, the ends of the rod in system S′ be at x'_1 and x'_2 in Fig.6.8.

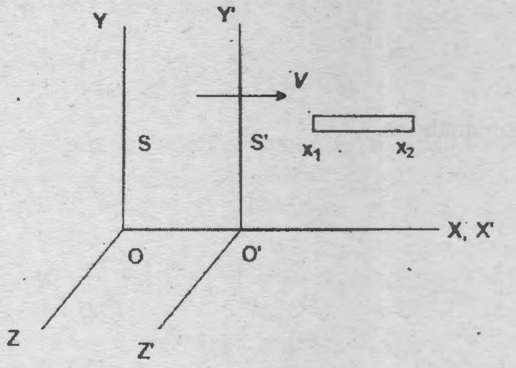


Fig. 6.8.

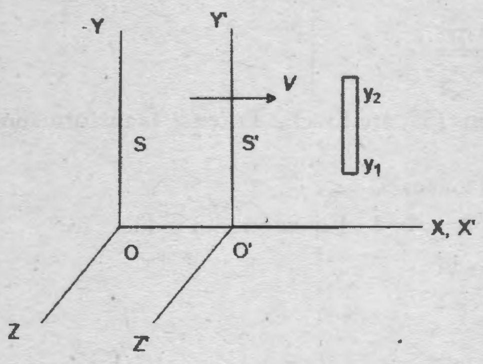


Fig. 6.9.

Length of the rod in system S′ is, $(x'_2 - x'_1) = L_0$ (33)

Here L_0 is the proper length.

Proper length may be defined as the length when measured in a frame of reference in which the rod is stationary. We can determine the length of the rod from frame S also. Let x_1 and x_2 be co-ordinates of the ends of the rod at time t. Thus, the length of the rod (at rest in S′) observed by the observer from S is:

$$L = (x_2 - x_1) \quad (34)$$

Using Lorentz transformation equation (16), we have:

$$x'_1 = \lambda(x_1 - vt_1) \quad (35)$$

$$x'_2 = \lambda(x_2 - vt_2) \quad (36)$$

Where $\lambda = \frac{1}{\sqrt{1 - v^2/c^2}}$

Subtract (35) from (36) we get:

$$L_0 = x'_2 - x'_1 = \lambda(x_2 - vt_2) - [\lambda(x_1 - vt_1)]$$

$$= \lambda x_2 - \lambda vt_2 - \lambda x_1 + \lambda vt_1$$

$$= \lambda(x_2 - x_1) - \lambda vt_2 + \lambda vt_1$$

(Here $t_1 = t_2$) as the observer in S measures the positions of the two ends of the rod at same time t in S′)

We have thus,

$$x'_2 - x'_1 = \lambda(x_2 - x_1)$$

Or $L_0 = \lambda L$

$$L = \frac{L_0}{\lambda} = L_0\sqrt{1-\frac{v^2}{c^2}} \qquad \text{Length contraction} \qquad (37)$$

$L < L_0$ $\qquad$ [for $\sqrt{1-\frac{v^2}{c^2}} << 1$]

This implies if L_0 the proper length of the rod at rest is, it will be measured as contracted to $L_0\sqrt{1-\frac{v^2}{c^2}}$ by an observer in S'.

If we consider the case in which the length of rod is perpendicular to the direction of frame (along Y axis) as shown in Fig.6.9. and if y_1 and y_2 are the co-ordinates of ends of rod relative to frame S corresponding to co-ordinates y_1' and y_2' in frames S', then:

Proper length: - $L_{0y} = y_2' - y_1'$ in S'

Length of rod in frame S,

$L_y = y_2 - y_1$

From Lorentz transformation, we have, $y = y'$

So $y_1 = y_1'$ and $y_2 = y_2'$

$L_{oy} = y_2' - y_1'$

$= y_2 - y_1 = L_y$

This implies that the length of the rod does not change in a direction perpendicular to the direction of motion.

In nutshell, the length of the moving rod is contracted along the direction of motion by a factor $\sqrt{1-v^2/c^2}$ while there is no contraction along a direction perpendicular to the direction of motion. This is known as Lorentz Fitzergerald contraction.

The above said phenomenon is not the result of some force "squeezing" the rod but is a real physical phenomenon with observable physical effects.

6.11.2 Length Contraction Evidence

(a) Muon Decay

The presence of muons near the surface of the earth is the most important evidence for length contraction. Muons are created in the upper region of the atmosphere at height of about 10 km by cosmic ray particles arriving from outer space. Muons travel at a speed of 0.998c, are highly unstable with mean life of 2.2×10^{-6} sec. Before they decay, they travel a distance of, $L = 0.998 \times 3 \times 10^8 \times 2.2 \times 10^{-6} = 660$m. But, how is it possible of their travelling distance 10 km to reach the surface of earth. By considering relativistic length contraction, this puzzle can be solved.

The above distance L calculated is the distance which a muon travel in its own frame of reference which a muon travel in its own frame of reference. If L_0 is the distance travelled in earth's frame then:

$$L_0 = \frac{L}{\sqrt{1-v^2/c^2}} = \frac{660}{\sqrt{1-(0.998)^2}} = 10.5 \text{ km}$$

This implies that irrespective of their shorter lifetime, muons can reach earth surface from high altitudes where they are created.

6.11.3 Time Dilation

Consider two frames of reference S and S' such that S is at rest and S' is moving along +x direction with velocity v relative to S. Let a clock be placed in frame S which is at rest and gives a signal at time t_1 in system S and corresponding in S' frame when measured by observer in S' frame.

In frame S, the time interval between the two events (it may be light flashes or ticks of clock) is given as: t_1'

$$\Delta t = t_2 - t_1 \tag{38}$$

In frame S' or moving frame, it is equal to $\Delta t' = t_2' - t_1'$ (39)

From Lorentz transformation,

$$t_1' = \frac{t_1 - \frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad t_2' = \frac{t_2 - \frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \tag{40}$$

Putting (40) in (38) i.e.

$$\Delta t' = t_2' - t_1'$$

$$= \frac{(t_2 - t_1)}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\Delta t' = \gamma\Delta t \qquad \text{Where } \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \tag{41}$$

i.e. $\Delta t' > \Delta t$

Equation (41) shows the dilated or increased time interval for a moving observer. This is called time dilation. In other words, "moving clock appears to go slow or it can be also expressed in a way that "every clock appears to go at its fastest rate at rest frame, relative to the observer in moving frame. e.g. If the time interval between two events in the spacecraft is $\Delta t'$ as observed in the spacecraft then a longer time interval at between two events can be observed by observer on earth.

In an inertial frame, if the two events occur simultaneously at the same point, then the time interval between them measured in that frame, is called proper time interval. If the same time interval is measured by an observer in any other inertial frame, it is always greater. This is known as time dilation.

For example, let us consider a stationary frame S' (a railway station). Consider an inertial frame S fixed in a train moving with a constant speed relative to a station in Fig. 6.10.

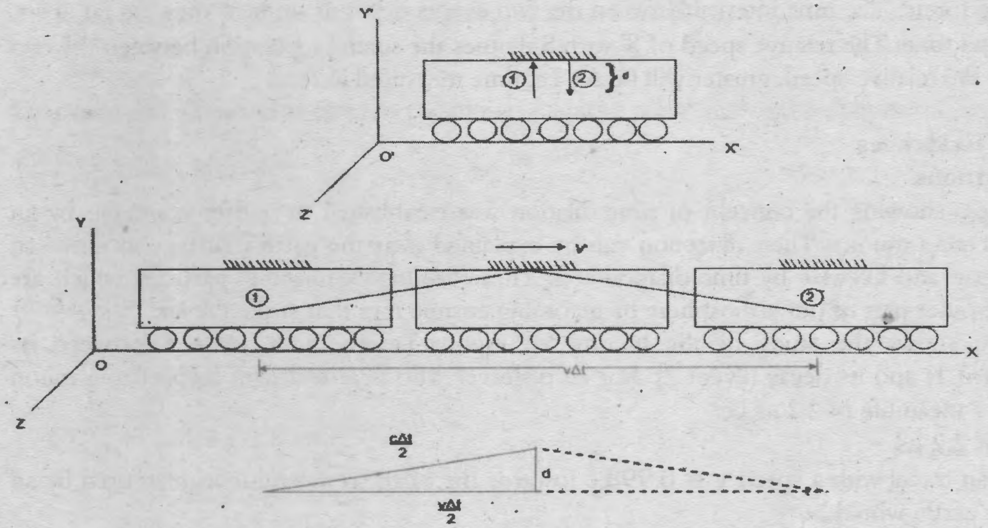


Fig. 6.10.

Think of the case when an observer in rest frame sends a flash light or light pulse to the mirror which is fitted at the ceiling with speed c through some source (event1). On reflection it returns back to the observer at the starting point (event 2). The time measured by a stationary observer in S' frame is Δt_0, the time interval between the two events as measured by observer in S' is $\Delta t_0 = \frac{2d}{c} \Delta t_0$

Where d is the distance between the source of light pulse and mirror at the ceiling.

In the stationary frame S since the mirror is moving to the right with speed v when light pulse strikes the mirror it moves a distance $v\left(\frac{\Delta t}{2}\right)$ where Δt is time taken by light pulse for its round trip as measured in S frame.

We get then,

$$\left(\frac{c\Delta t}{2}\right)^2 = \left(\frac{v\Delta t}{2}\right)^2 + d^2$$

$$(c^2 - v^2)(\Delta t)^2 = (2d)^2$$

$$\Delta t = \frac{2d}{\sqrt{c^2 - v^2}} = \frac{2d/c}{\sqrt{1 - v^2/c^2}}$$

But $\frac{2d}{c} = \Delta t'$

$$\Delta t = \frac{\Delta t'}{\sqrt{1 - v^2/c^2}}$$

or $T = \frac{T_0}{\sqrt{1 - v^2/c^2}}$

The above relation shows,

$\Delta T > \Delta T_0$

This implies that in S' frame, the occurrence of the two events takes place at the same location in space although in S frame, this happens at different locations. Thus, a stationary clock will measure as longer time interval between events occurring in a moving frame of reference in comparison to clock in moving frame. The time interval between the two events depends on how they are far apart, in both space and time. The relative speed of S′ with S defines the spatial separation between the two events. Greater the relative speed, greater will be ΔT i.e. time measured in S.

Experimental Evidences

(a)Decaying Muons

The first example showing the concept of time dilation was established over fifty years ago by an experiment detecting muons. Their detection can be explained near the earth's surface in terms of length contraction and likewise by time dilation too. These are highly unstable particles which are produced at the outer part of our atmosphere by incoming cosmic rays that strike the air.

The experiment shows the study of the lifetime of muons i.e. the time interval between its production (event 1) and its decay (event 2). For an observer who is at rest with respect to a muon would measure a mean life of 2.2 μs i.e.

$$\Delta T_o = 2.2\ \mu s$$

These muons can travel with a speed $v = 0.998\ c$ towards the earth so mean life as measured by an observer on the earth, would be

$$\Delta T = \frac{\Delta T_0'}{\sqrt{1 - v^2/c^2}} = \frac{2.2 \times 10^{-6}}{\sqrt{1-(0.998)^2}} = 34.8\ \times\ 10^{-6}\ \text{sec}$$

$$= 34.8\ \mu\text{sec}$$

Which is around 16 times more than ΔT_0

The distance travelled by muon in a time interval of 34.8 μsec would be: -

$$vt = 2.994\ \times 10^8 \times 34.8 \times 10^{-6}$$

$$= 1.04 \times 10^4 m = 10.4\ km$$

(b)Decay of pions

Pions are produced as a result of collision between protons and certain nuclei. These are unstable particles with a half-life of $1.8\ \times 10^{-8}$ sec i. e

$$T_{1/2} = 1.8 \times 10^{-8}\ \text{sec} \tag{42}$$

These when produced are taken down through an evacuated pipe to the site of experiment and travel with a speed of $v = 0.999c$. The distance travelled by them is about $L = 1km$

As measured in laboratory, time taken by the pions to reach the experimental site is:

$$T \approx \frac{L}{v} \approx \frac{10^3}{3 \times 10^8} = 3.3\ \mu\text{sec} \tag{43}$$

from (42) and (43), $\quad \frac{T}{T_{1/2}} = \frac{3.3 \times 10^{-6}}{1.8 \times 10^{-8}} = 183$

i.e., $T\ \approx 183\ T_{1/2}$ is the time taken by pions to reach the site or, 183 half – lives.

Fraction of pions that remain undecayed during its journey is:

$$\left(\frac{1}{2}\right)^n = \left(\frac{1}{2}\right)^{183} = 8.2\ \times 10^{-56}$$

No pions will reach the experiment site but this actually does not take place looking the example relativistically.

Consider the two frames that refer to T and $T_{1/2}$ separately. Let T_{lab} be the time of flight of pions as measured in laboratory frame:

i.e., $T_{lab} = 3.3 \times 10^{-6}$ sec (44)

$T_{1/2}$ is measured in frame fixed to the pions i.e., pions are at rest in this frame

$$T_{1/2}^{rest} = 1.8 \times 10^{-8} \text{ sec} \quad (45)$$

Thus, $\Delta T = \frac{T_0}{\sqrt{1 - v^2/c^2}}$

or $\Delta T_{1/2}^{lab} = \gamma \ \ T_{1/2}^{rest}$

$$= \frac{1}{\sqrt{1-(0.99999)^2}} \times 1.8 \times 10^{-8} \text{ sec} \quad (46)$$

$$\Delta T_{1/2}^{lab} = 1000 \times 1.8 \times 10^{-8} \text{ sec} \quad (47)$$

$$= 1.8 \times 10^{-5}\text{sec}$$

from (44) and (47), we have;

$$T_{lab} \approx 0.2\, T_{1/2}^{lab} \quad (48)$$

Only one – fifth of the half – life do the pions flight in the pipe. Only few of the pions decay during this time and hence they reach to the experimental site. This proves the time dilation of the pions in the physics laboratories due to the theory of relativity.

The Twin Paradox (Consequences of Length Contraction and Time Dilation)

Consider two twins T_1 and T_2 each with 30 years of age. Let T_1 remains at rest (S frame) on earth and T_2 travels within a space ship (S′ frame) with a velocity $\frac{\sqrt{3}}{2}$ c for 2 years along positive direction of X-axis and after 2 years he returns back to the same point he started.

According to T_2, his age is 30 +4 = 34 years but according to T_1, age of T_2 will be

$$\Delta T = \frac{T_0}{\sqrt{1 - v^2/c^2}} = \frac{4}{\sqrt{1-(\frac{\sqrt{3}c}{2c})^2}}$$

$$= \frac{4}{\sqrt{\frac{4-3}{4}}} = 4 \times 2 = 8 \text{ years}$$

i.e T_2' age will be 30 + 8 = 38 years

Similarly, age of T_1 according to T_1 would be 38 years whereas age of T_1 according to T_2 would then be 34 years older. Thus, this is a contradictory statement as both of them consider themselves to be younger than the other. How, then this paradox can be resolved? Of identical twins (T_1 and T_2), T_2 also takes journey into space and returns home to find T_1 has aged more.

If we can recall the Lorentz transformations are applicable to only inertial co-ordinate system i.e., for those moving at a constant velocity with respect to each other. Thus, T_1 who is at rest on earth is in an inertial co-ordinate system and can use the time dilation equation. The twin T_2 who returns back however is in a non-inertial co-ordinate system.

If T_2 is originally moving at a velocity v, to return back on earth then he has to decelerate his spaceship to zero velocity and then accelerate to –v velocity to travel back home. Thus, T_2 is not in an inertial co-ordinate system and is not subjected to the time dilation formula. On the other hand, the prediction made by twin T_1 who is on earth for him, time has slowed down for his twin T_2. When T_2 returns back he should find T_1 to be older than he is.

6.12 Relativity of Simultaneity

The events that occur at the same time are called simultaneous event. Consider two frames of reference S and S′. S′ moving relative to S with a constant velocity v along the direction of X. Let the two events occur simultaneously at points P_1 and P_2 in frame S.
The coordinate at P_1 be (x_1, y_1, z_1, t_1) and P_2 be (x_2, y_2, z_2, t_2) measured by an observer in frame S. Since the events are simultaneous in frame S, thus $t_1 = t_2$. Let $t_1'$and t_2' be the corresponding times of the same events in S′ frame, from Lorentz transformation equations:

$$t_1' = \frac{t_1 - \frac{vx_1}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \quad , \quad t_2' = \frac{t_2 - \frac{vx_2}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$t_2' - t_1' = \frac{t_2 - t_1}{\sqrt{1-v^2/c^2}} - \frac{\frac{v}{c^2}(x_2 - x_1)}{\sqrt{1-v^2/c^2}}$$

$$= \frac{v\,(x_2 - x_1)}{\sqrt{1-v^2/c^2}} \quad (\text{as } t_1 = t_2) \qquad (49)$$

For the events to be simultaneous in S′ frame, t_1' must be equal to t_2' or $t_2' - t_1'$ must be zero but this is not true as $x_1 \neq x_2$. Thus, the same two events are not simultaneous in S′ frame. Two events at different places P_1 and P_2 which are simultaneous for an observer at rest in frame S, are no longer simultaneous to an observer in frame S′ having linear motion relative to S′. In nutshell we can say, simultaneity is also relative, not absolute.

6.13 Velocity Addition (Transformations of Relative Velocities)

Consider two frames S and S′ moving with a relative velocity v with respect to S along +x direction. Let a particle P (x′ y′ z′ t′) where (x′ y′ z′ t′) are the space and time co-ordinates of P in frame S′ moves with velocity $\overrightarrow{u'}$ relative to frame S′ such that

$$\overrightarrow{u'} = \hat{\imath}\,\overrightarrow{u_x'} + \hat{\jmath}\,\overrightarrow{u_y'} + \hat{k}\overrightarrow{u_z'} \qquad (50)$$

Where $\overrightarrow{u_x'}$, $\overrightarrow{u_y'}$ and $\overrightarrow{u_z'}$ be components of $\overrightarrow{u'}$ along x, y, z axes respectively and $\hat{\imath}, \hat{\jmath}, \hat{k}$ be the unit vectors along x, y, z axis respectively.

The velocity components can be defined as:

$$u_x' = \frac{dx'}{dt'}, u_y' = \frac{dy'}{dt'}, u_z' = \frac{dz'}{dt'} \qquad (51)$$

Similarly, P (x y z t) are co-ordinates for particle in frame S and u its velocity relative to frame S s.t $\vec{u} = \hat{\imath}\overrightarrow{u_x} + \hat{\jmath}\overrightarrow{u_y} + \hat{k}\overrightarrow{u_z}$.

The velocity components along three axes are given as:

$$u_x = \frac{dx}{dt}, u_y = \frac{dy}{dt}, u_z = \frac{dz}{dt} \tag{52}$$

From inverse Lorentz transformations, we have

$$\left.\begin{aligned} x &= \frac{x' + v t'}{\sqrt{1 - \frac{v^2}{c^2}}} \\ y' &= y \\ z' &= z \\ t &= \frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned}\right\} \tag{53}$$

Taking differentials of above equations, we have

$$\left.\begin{aligned} dx &= \frac{dx' + v\,dt'}{\sqrt{1 - \frac{v^2}{c^2}}} && \text{(i)} \\ dy &= dy' && \text{(ii)} \\ dz &= dz' && \text{(iii)} \\ dt &= \frac{dt' + v\frac{dx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} && \text{(iv)} \end{aligned}\right\} \tag{54}$$

Dividing 54(i) by 54(iv) we get

$$u_x = \frac{dx}{dt} = \frac{dx' + v\,dt'}{dt' + \frac{v\,dx'}{c^2}} = \frac{\frac{dx'}{dt'} + v\frac{dt'}{dt'}}{\frac{dt'}{dt'} + \frac{v}{c^2}\frac{dx'}{dt'}}$$

(Divide numerator and denominator by dt′)

Or $u_x = \dfrac{u'_x + v}{1 + \frac{v}{c^2}u'_x}$ [from equation (51): $\dfrac{dx'}{dt'} = u'_x$] (55)

Similarly, from 54 (ii) and 54 (iv), we get:

$$u_y = \frac{dy}{dt} = \frac{dy'\sqrt{1 - \frac{v^2}{c^2}}}{dt' + \frac{v\,dx'}{c^2}} = \frac{\frac{dy'}{dt'}\sqrt{1 - \frac{v^2}{c^2}}}{\frac{dt'}{dt'} + \frac{v}{c^2}\frac{dx'}{dt'}} = \frac{u'_y\sqrt{1 - \frac{v^2}{c^2}}}{1 + \frac{v}{c^2}u'_x}$$

Or $$u_y = \frac{u'_y\sqrt{1 - \frac{v^2}{c^2}}}{1 + \frac{v}{c^2}u'_x} \tag{56}$$

From equations 54(iii) and 55(iv), we get

$$u_z = \frac{u'_z\sqrt{1 - \frac{v^2}{c^2}}}{1 + \frac{v}{c^2}u_x} \tag{57}$$

Equation (55), (56), (57) thus represents the components of u_x, u_y, u_z of velocity $\vec{u}$ in frame S respectively. If the particle is moving along x axis i.e in the direction of v in frame S' with velocity u', then if $u'_x = u'$ and $u_x = u$

From equation (55), we have

$$u_x = \frac{u'+v}{1+\frac{u'v}{c^2}} \qquad (58)$$

Thus equation (55), (56), (57), (58) all represents the relativistic law of addition of velocities. However, they can be reduced to Galilean results or using classical mechanics where u' and v are small in comparison to the velocity of light i.e. $\frac{u'v}{c^2} << 1$, then

$$u = u' + v$$

If $u' = c$ i.e consider a particle photon travelling with velocity c along the direction of x axis, then from equation (58),

$$u_x = \frac{c+v}{1+\frac{cv}{c^2}} = \frac{c+v}{\frac{c+v}{c}} = c \qquad (59)$$

A particle moving with velocity of light c will be observed c in all inertial frames whatever their relative speed is, the "velocity of light is an absolute constant independent of the motion of the reference system".

If $u' = c$ and $v = c$, then from equation (58) we have:

$$u_x = \frac{c+c}{1+\frac{c^2}{c^2}} = \frac{2c}{2} = c \qquad (60)$$

or $u_x = c$

"The addition of any velocity to the velocity of light merely reproduces the velocity of light". Thus, velocity of light is the maximum attainable velocity, we can found in nature.

6.14. Variation of Mass with Velocity

According to the classical concept, the mass of a moving body is equal to that of stationary body but according to Einstein theory, the mass of a moving body changes with velocity and is different from the mass at rest. Thus, like length and time, mass also depends on the velocity of the moving object. Consider two frames of reference S and S', S' moving with a velocity v relative to frame as shown in Fig.6.11.

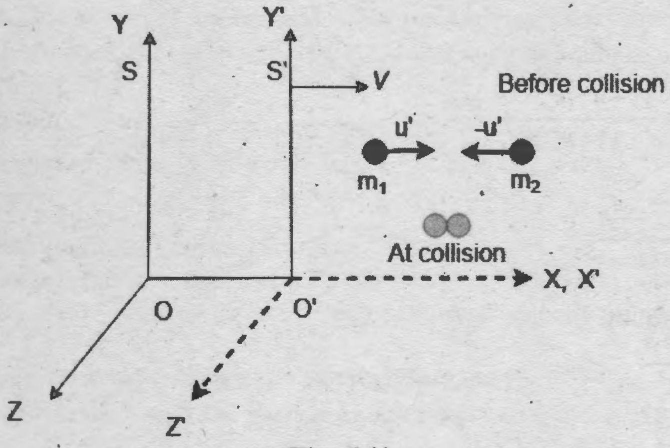


Fig. 6.11.

Let us take two particles with masses in frame S′ travelling with equal and opposite velocities u′ and – u along +x direction along x axis. The two particles collide and after impact coalesce into one particle. By the law of conservation of momentum, this coalesced particle is at rest i.e.

Momentum of mass m_1 + momentum of mass m_2 = momentum of coalesced body i.e.

mu + (- mu) = 0 (€oalesced body at rest in frame S′) (61)

Also, according to the principle of conservation of mass, their combined mass would be 2m (m + m) in frame S′ after impact.

On contrary for an observer in frame S, is observing this impact in a different way. If m_1 and m_2 be the masses of the two particles seen from frame S and if the velocities of the two particles u_1 and u_2 in frame S corresponds to u' and $-u'$ in frame S′, then

$$u_1 = \frac{u'+v}{1+u'v'c^2}, \qquad u_2 = \frac{-u'+v}{1-u'/v/c^2} \tag{62}$$

If m_1 is travelling with velocity u_1 and m_2 with u_2, then according to the principle of conservation of momentum;

$$m_1u_1 + m_2u_2 = (m_1 + m_2)v \tag{63}$$

From equation (62), put values of u_1 and u_2 in equation (63), we get

$$m_1\left(\frac{u'+v}{1+\frac{u'v}{c^2}}\right) + m_2\left(\frac{-u'+v}{1-\frac{u'v}{c^2}}\right) = (m_1 + m_2)v$$

$$m_1\left(\frac{u'+v}{1+\frac{u'v}{c^2}} - v\right) = m_2\left(v - \frac{-u'+v}{1-\frac{u'v}{c^2}}\right)$$

$$\text{or } m_1\left[\frac{u'+v-v-\frac{u'v^2}{c^2}}{1+\frac{u'v}{c^2}}\right] = m_2\left[\frac{v-\frac{u'v^2}{c^2}+u'-v}{1-\frac{u'v}{c^2}}\right]$$

not infinite, as photon has rest mass as zero and undefined moving mass. Ultimately, we can say, no material particle can have velocity equal to or greater than velocity of light.

Special Cases

(1) If $v << c$, $\frac{v^2}{c^2}$ is negligible in equation (69) compared to 1. So $m = m_0$

(2) If $v \approx c$, $\sqrt{1 - v^2/c^2}$ is < 1, $m > m_0$

(3) If $v = c$, body possess infinite/ undefined mass.

In particle accelerators, effective mass of particles has been verified experimentally for electrons and protons. Their velocities were subjected to the velocities close to velocity of light.

6.15 Relativistic Momentum and Newton's Second Law

Relativistic momentum can be given for a particle of rest mass m_0, moving with a velocity v as

$$p = mv = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad \text{(Using equation (69)}$$

$$= \gamma m_0 v \tag{70}$$

The law of conservation of momentum, is applicable to relativity theory also.

The relativistic generalization of Newton's second law is given as:

$$F = \frac{dp}{dt} = \frac{d}{dt}(mv) = \frac{d}{dt}\frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{71}$$

$$\frac{d}{dt}(mv) = m\frac{dv}{dt} + v\frac{dm}{dt} = ma + v\frac{dm}{dt}$$

Since there is dependence of m on v, so $\frac{dm}{dt}$ does not vanish if the velocity of particle varies with time. Thus, Newton's law, $F = ma$ is invalid in relativity theory, if mass variation is not taken into consideration and Lorentz transformation equations are not applied.

6.16 Relativistic Kinetic Energy

For a body of mass m moving with velocity v the kinetic energy can be given as $\frac{1}{2}mv^2$ in Newtonian Mechanics. Consider a body at rest which is acted under a force so that it moves. The work – energy principle defines the kinetic energy of a body as the amount of work done by the force to move the body from its state of rest to the state of motion with velocity v. If F be the force applied on a body such that it causes the body to move a distance dx along its direction, then the work is done by the force and the kinetic energy of the body increases:

$$dk = Fdx \tag{72}$$

From Newton's second law,

$$F = \frac{dP}{dt} = \frac{d}{dt}(mv)$$

Since both m and v are variables

$$F = m\frac{dv}{dt} + v\frac{dm}{dt}$$

Substitute in equation (72) above

$$dk = \left(m\frac{dv}{dt} + v\frac{dm}{dt}\right)dx$$

$$= m\left(\frac{dx}{dt}\right)dv + v\left(\frac{dx}{dt}\right)dm$$

$$= mvdv + v^2dm \qquad (73)$$

From the relativistic mass variation relation with velocity, we have

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$$

$$\text{or } m^2 = \frac{m_0{}^2c^2}{c^2 - v^2}$$

$$\text{or} \qquad m^2c^2 - m^2v^2 = m_o{}^2c^2$$

On differentiation,

$$(2m\,dm)c^2 - (2m\,dm)v^2 - (2v\,dv)m^2 = 0$$

$$\text{or } m\,v\,dv + v^2dm = c^2dm \qquad (74)$$

Comparing (73) and (74)

$$dk = c^2\,dm \qquad (75)$$

For a body which is accelerated from rest to a velocity v, mass increases from m_0 to m. If K is the kinetic energy acquired by the body, then:

$$\int_o^k dk = c^2\int_{m_o}^{m} dm$$

$$\text{Or } K = c^2(m - m_0) = c^2\left(\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} - m_0\right)$$

$$K = m_0c^2\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right) \qquad (76)$$

Equation (76) represents the relativistic expression for kinetic energy of a body

Special Cases: -

(1) If v << c, equation (76) must reduce to Newtonian expression. Let us expand γ binomially and retain to small powers of $\frac{v}{c}$, as it is very small.

$$\text{Thus,} \quad \gamma = \left(\frac{v^2}{c^2}\right)^{\frac{-1}{2}} = 1 + \frac{1}{2}\frac{v^2}{c^2}$$

Substitute in equation (76)

$$K = m_0c^2(1 + \frac{1}{2}\frac{v^2}{c^2} - 1) = \frac{1}{2} m_0v^2$$

$K = \frac{1}{2} m_0v^2$, (Same as Newtonian expression)

(2). If $v \rightarrow c$, $k \rightarrow \infty$, infinite amount of work would be required to be done on the particle so as to accelerate it to the speed of light. So, c plays the limiting speed for material particles.

Rewriting equation (76) we get:-

$K = (m - m_0)c^2$

or $mc^2 = K + m_0c^2$ (77)

Where mc^2 = total energy of the body,

K = kinetic energy of the body, is an increase in the mass of body due to relative motion,

m_0c^2 = rest mass energy of body which is the energy possessed by body in a reference frame which is at rest. It may also be considered as internal energy of the body.

Total energy is:

$$E = mc^2 = \frac{m_0c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma\, m_0c^2 \quad (78)$$

Rest mass energy is: $E_0 = m_0c^2$ (79)

Thus, Total energy, E = rest mass energy + relativistic kinetic energy

$E = m_0c^2 + (m - m_0)c^2 = mc^2$

$E = mc^2$ (80)

Equation (80) represents, Einstein mass energy relation which states universal equivalence between mass and energy. It explains us the interchangeability of mass and energy and are considered as single entity. The conservation laws are separate for mass and energy in classical physics although they combine into a single conservation law "the law of conservation of mass-energy." Mass can be either created or destroyed, for this an equivalent amount of energy vanishes or gets created and vice a versa. Therefore, a mass m is equivalent to the energy $E = mc^2$, where c is the velocity of light in vacuum. E can be presented in various forms like kinetic energy, heat energy, a light or other form of energy. Huge amount of energy will be released on annihilation of matter due to the conversion factor c^2 e.g if 1 g of matter is annihilated an equivalent of 9×10^{13}J of energy is liberated.

6.16.1 Experimental Evidence for Mass-Energy Relation

The most important direct evidences of mass-energy equivalence are: -

(1)During nuclear fusion process of two neutrons and two protons to form $_2He^4$ nucleus, an enormous amount of energy is released

$2\,_1H^1 + 2\,_0n^1 \rightarrow\, _2He^4$

Mass of reactants = mass of 2 protons + mass of 2 neutrons

Mass of 2 protons = $2\,_1H^1$

= 2 × mass of hydrogen nucleus

= 2 × 1.00815 a.m.u

= 2.01630 a.m.u

Mass of 2 neutrons = 2 × mass of neutrons

$= 2 \times 1.00898$ a.m.u
$= 2.01796$ a.m.u
Mass of reactants = $(2.01630 + 2.01796)$ a.m.u
$= 4.03426$ a.m.u
Mass of product = mass of helium nucleus
= mass of $_2He^4$
$= 4.00387$ a.m.u
Decrease in mass or mass defect = $(4.03426 - 4.00387)$ a.m.u
$= 0.03039$ a.m.u
$= 0.03039 \times 931.1$ (1a.m.u = 931.1MeV)
$= 28.3$ MeV

Thus, a nuclear fusion releases an energy of 28.3 MeV. This forms the basis of solar energy.

(2) In nuclear fission reaction, the heavy nuclei of certain materials like uranium may form fission nuclei. The total mass of these is lesser in comparison to the parent uranium nuclei. The difference in rest mass Δm appears in the form of equivalent energy $\Delta E = \Delta m.c^2$. The production of enormous amount of energy during fission reaction is the principle of atom bomb.

(3) When high energy radiation, with energy greater than $2m_0c,^2$ m_0 being the rest mass of electron (positron), passes near a heavy nucleus, an electron position pair is created wherein energy is converted into mass. This is known as pair production.

$2h\nu \rightarrow {}_{+1}e^0 + {}_{-1}e^0$

However, when an electron collides with a positron both the particles disappear giving rise to two photons. This is known as annihilation of matter.

${}_{+1}e^0 + {}_{-1}e^0 = 2h\nu$

6.16.2 Binding Energy

The nucleus of an atom consists of protons and neutrons bounded together by strong attractive forces. In order to break nucleus into its constituents a large amount of energy must be supplied. This energy is called binding energy of a nucleus and has an equivalent mass, $\Delta m = \frac{\Delta E}{c^2}$

6.16.3 Mass Defect: Δm

When the original atomic mass is less than the sum of individual masses of its constituent particles, this difference in mass is known as mass defect and is given as: -

$\Delta m = (m_n + m_p) - m_0$

$(m_n + m_p)$ = total mass of neutron and proton

m_0 = mass of original atom

Now, if a nucleus has atomic number Z and atomic weight Δ, then:

$\Delta m = Z.m_H + (A - Z)_{m_n} - M_{Z,A}$

Where m_H is the mass of hydrogen atom, m_n mass of neutron and $M_{Z,A}$ observed mass of atom under consideration e.g., deuteron with Z = 1, A = Z

Mass of deutron, $M_{1,2} = 3.3442 \times 10^{-27}$kg

$m_H = 1.6734 \times 10^{-27}$kg

$m_n = 1.6748 \times 10^{-27}kg$

Mass defect of deuteron,

$(\Delta m)_{H_2} = 1m_H + 1m_n - M_{12} \approx 3.98 \times 10^{-30}kg$

So, equivalent energy, $\Delta E = \Delta m.c^2$

$= 3.98 \times 10^{-30} \times (3 \times 10^8)^2$

$= 2.238$ MeV

Which is in agreement with the measured experimental value 2.226 MeV of binding energy of deuteron nucleus.

6.17 Relativistic Energy-Momentum

Relativistic momentum of a particle of mass m_0 moving with a velocity v is given by

$$p = \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} \tag{81}$$

Relativistic energy is given as:

$$E = \frac{m_0 c^2}{\sqrt{1-\frac{v^2}{c^2}}} \tag{82}$$

Squaring equation (81) and multiplying by c^2

$$p^2c^2 = \frac{m_0^2 v^2 c^2}{1 - v^2/c^2} \tag{83}$$

Squaring equation (82) we get

$$E^2 = \frac{m_0^2 c^4}{1 - v^2/c^2} \tag{84}$$

Subtracting from (84) from (83)

$$E^2 - p^2c^2 = \frac{m_0^2c^4 - m_0^2v^2c^2}{1-\frac{v^2}{c^2}} = m_0^2c^4$$

$$\text{or } E^2 = m_0^2c^4 + p^2c^2 \tag{85}$$

$$\text{or } E = \sqrt{m_0^2c^4 + p^2c^2} = \sqrt{E_0^2 + p^2c^2}$$

$$E = \sqrt{E_0^2 + p^2c^2} \tag{86}$$

6.17.1 Physical Significance of Relativistic Energy-Momentum Relation

It can be seen from equation (85), the total energy E, rest mass energy m_0c^2 of a relativistic particle and energy of the particle equivalent to its photonic energy pc are related to each other with right angled triangle as shown in the Fig.6.12. Again, it seems that the total energy of a relativistic particle is a complex quantity i.e., $E = m_0c^2 + ipc$. It means rest mass energy and the equivalent photonic energy of a relativistic particle cannot be added algebraically.

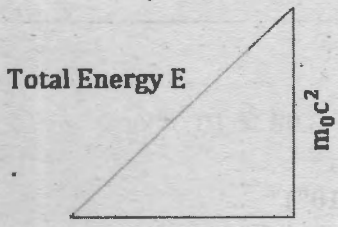


Fig. 6.12.

6.17.2 Massless Particles

In classical physics, for a particle to have momentum and energy, it must have non-zero rest mass. But equation (85), shows that a particle can have momentum as well as energy even when the rest mass is zero.

So $E = pc$ (for massless particle) (87)

Massless particles travel with speed of light e.g., photons and neutrino.

Equation (81) and (82) above shows that massless particles with a speed less than speed of light cannot exist. If $m_0 = 0$ and $v < c$, $p = E = 0$ i.e., such kind of particles have neither momentum nor energy and hence it does not exist. But if v = c then p and E are indeterminate and can have any value.

6.17.3 Transformation of Maxwell's Equations Under Lorentz Transformations

Here are the Maxwell's equations of electromagnetism in integral from;

(i) $\oint \mathbf{E}.\,d\mathbf{S} = \frac{q}{\epsilon_0}$ Gauss's Law of Electrostatics

(ii) $\oint \mathbf{B}.\,d\mathbf{S} = 0$ Gauss's Law of Magnetostatics

(iii) $\oint \mathbf{E}.\,d\mathbf{l} = -\frac{\partial \phi_B}{\partial t}$ Faraday's Law of EM Induction

(iv) $\oint \mathbf{H}.\,d\mathbf{l} = I + I_d = \oint (J + J_d).\,dS$ Modified Ampere's Circuital Law

And in differential form;

$$\nabla.\mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla.\mathbf{B} = 0$$

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

$$\nabla \times \mathbf{H} = \mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

We know that the invariance of any physical law/equation, under Galilean or Lorentz transformations is that, no change in the law/equation, if the system is put under non-relativistic (Galilean) or Relativistic (Lorentz) motion. Even the form or physical variable involved in the equation/ law should not be changed for the condition of invariance.

Now in the case of invariance of Maxwell's equations of electromagnetics, we should keep in mind that one of equations is the equation of electrostatics, i.e., it is written for statics charge and other is written for statics magnetic field. Now, it is well known fact that as soon as the charges as well as magnetic field set into motion, some other physical quantities come into picture. For example, as soon as the frame containing charge at rest, set into constant motion, the current and magnetic field also come into picture. And if a conductor having some charge density, set into motion, the charge density will also change in accordance with the Lorentz's length contraction concept. It means that Maxwell's equations must be modified in moving frame of reference with some other physical variables. The equations (3) and (4) should be affected greatly due to the "time dilation" concept of Lorentz. So, from above discussion, it is clear that Maxwell's equations may not seem to be invariant, in as it is form. It means the analogy of Maxwell's equations for moving frame of reference, must be modified for the proper invariance under Lorentz's transformation.

Here we will prove the invariance of Maxwell's equations in differential form, as they are related to the explanations of more new physical phenomena like electromagnetic waves, etc.

Here, we have the transformation of some physical parameters i.e., charge density ρ, electric field and magnetic field under Lorentz transformations.

As, we have transformation of length from one frame of reference w.r.t. another, in case of relativistic motion. Using this concept, the charge density of a conductor can be transformed from one frame w.r.t. relativistically moving frame. If a charged conductor has proper charge density ρ_0 then in accordance with the length contraction concept of Lorentz, the relativistic charge density ρ is written as;

$$\rho = \frac{\rho_0}{\sqrt{1-\frac{u^2}{c^2}}} \qquad \text{(Charge density dilation equation)} \qquad (88)$$

Where u is the velocity of charged conductor under consideration relative to the reference frame, from which ρ is observed. If ρ_0 is the static charge density then obviously, Maxwell's first equation for a frame in relative motion will be;

$$\nabla . D = \frac{\rho_0}{\sqrt{1-\frac{u^2}{c^2}}} \qquad (89)$$

Now, keeping special theory of relativity into consideration, we have the Lorentz transformation equations for electric and magnetic field in a moving frame w.r.t. stationary frame as;

$$\left.\begin{array}{lll} E'_x = E_x \text{ (i)} & E'_y = (E_x - uB_z) \text{ (ii)} & E'_z = \gamma(E_z + uB_y) \text{ (iii)} \\ B'_x = B_x \text{ (iv)} & B'_y = (B_y + uE_z) \text{ (v)} & B'_z = \gamma(B_z - uE_y) \text{ (vi)} \end{array}\right\} \qquad (90)$$

Here we should also keep in mind that Galilean transformation transform the three space coordinates only and the time is kept constant. That means, the physical variables can be defined in three coordinate system. But Lorentz transformation contains fourth coordinate, i.e., time coordinate. So, under Lorentz transformation, the transformation of any physical quantities must involve four-vector form, i.e., three space-coordinate and fourth time-coordinate. It means, without involving time coordinate, Maxwell's equations should not be transformed from one frame of reference w.r.t. another frame. It is very tedious, complicated and time consuming, to transform all four equations in moving frame w.r.t. a frame at rest and then showing the invariance. So first we will convert Maxwell's four equations into two, and then go for transformation and will show invariance under Lorentz transformations.

6.17.4 Maxwell's Equations in the Form of Electromagnetic Potential

As we know, the electric field potential 'ϕ' is related to electrostatic field and magnetic potential '**A**', is related to magnetic field. But as soon as the study of electromagnetism come into picture, '**A**' and 'ϕ' will no longer exist as separate physical quantity, rather they become a single physical quantity, called electromagnetic potential, written as $(\mathbf{A}, \phi)$. As magnetic potential is a vector quantity having $\mathbf{A_1}$, $\mathbf{A_2}$ and $\mathbf{A_3}$ components and electric potential is a scalar, so in classical electrodynamics, the electromagnetic potential is called, Electromagnetic Four Potential. In classical electrodynamics, (A_1, A_2, A_3, ϕ), seems to be four separate physical quantity. But in case of relativistic electrodynamics, the electromagnetic potential can be written as a single variable $\mathbf{A_\mu}$, having four vectors $(\mathbf{A_1, A_2, A_3, A_4})$, called as electromagnetic four-vector potential and written as;

$$\mathbf{A_\mu} = \left(\mathbf{A}, \frac{i\phi}{c}\right) \text{ or } (\mathbf{A_1, A_2, A_3, A_4}). \qquad (91)$$

It can easily be proved that $\frac{i\phi}{c} = \mathbf{A_4}$, becomes the fourth vector of electromagnetic potential.

Now, we will show that Maxwell's four equations of electromagnetism can be transformed into two equations in four-vector potential from.

We have Maxwell's four equations of electromagnetic field as;

$$\nabla.\mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla.\mathbf{B} = 0$$

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

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\frac{\partial\mathbf{E}}{\partial t}$$

In classical electrodynamics, the electric and magnetic fields vectors can also be expressed as

$$\mathbf{B} = \nabla\times\mathbf{A} \Rightarrow \nabla.\mathbf{B} = \nabla.(\nabla\times\mathbf{A})$$

$$\Rightarrow \nabla.\mathbf{B} = 0$$

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t} \Rightarrow \nabla\times\mathbf{E} = -(\nabla\times\nabla)\phi - \frac{\partial(\nabla\times\mathbf{A})}{\partial t}$$

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

Now using first equation

$$\nabla.\mathbf{E} = \frac{\rho}{\epsilon_0}$$

$$\nabla.\left(-\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}\right) = \frac{\rho}{\epsilon_0}$$

$$\nabla^2\phi + \frac{\partial(\nabla.\mathbf{A})}{\partial t} = -\frac{\rho}{\epsilon_0} \qquad (92)$$

Now using fourth equation

$$\nabla\times\nabla\times\mathbf{A} = \mu_0\mathbf{J} + \mu_0\epsilon_0\frac{\partial}{\partial t}\left(-\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}\right)$$

$$\nabla(\nabla.\mathbf{A}) - \nabla^2\mathbf{A} = \mu_0\mathbf{J} + \mu_0\epsilon_0\left(-\nabla\frac{\partial\phi}{\partial t} - \frac{\partial^2\mathbf{A}}{\partial t^2}\right)$$

On rearranging and putting $\mu_0\epsilon_0 = \frac{1}{c^2}$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla\left(\nabla.\mathbf{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) = -\mu_0\mathbf{J}$$

Here, $\nabla.\mathbf{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} = 0$ (Lorentz condition),

so

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J} \qquad (93)$$

Now putting $\nabla.\mathbf{A} = -\frac{1}{c^2}\frac{\partial \phi}{\partial t}$, from Lorentz equation into (I), we get

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \qquad (94)$$

Equations (93) and (94) are equivalent Maxwell's equations of electromagnetic field for the case of classical electrodynamics.

Now the equation (93) and (94) can also be written in d'Alembert operator form, $\Box^2$. Briefly, it can be easily understood in following way;

As $\nabla = \left(\hat{\imath}\frac{\partial}{\partial x} + \hat{\jmath}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)$ is a three-dimensional vector operator

$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is Laplacian operator.

In the same manner $\Box^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial_2}{\partial t^2} = \nabla^2 - \frac{1}{c^2}\frac{\partial_2}{\partial t^2}$, is called d' Alembert four-dimensional operator.

So, in terms of d' Alembert operator equations (93) and (94) can be written as;

$$\Box^2 \mathbf{A} = -\mu_0 \mathbf{J} \qquad (95)$$

$$\Box^2 \phi = -\frac{\rho}{\epsilon_0} \qquad (96)$$

Equations (95) and (96) are the equivalent form of Maxwell's equations of electromagnetism using d'Alembert. But using relativistic electrodynamics, where electromagnetic potential is written as electromagnetic four-vector potential, $\mathbf{A}_\mu = (\mathbf{A}, \frac{i\phi}{c})$, the Maxwell's four equations of electromagnetism can be converted into single one i.e.;

$$\Box^2 \mathbf{A}_\mu = -\mu_0 \mathbf{J}_\mu$$

Where $\mathbf{A}_\mu = (A_1, A_2, A_3, A_4) = (\mathbf{A}, \frac{i\phi}{c})$ and $\mathbf{J}_\mu = (J_1, J_2, J_3, J_4) = (\mathbf{J}, ic\rho)$ are the four-vector potential and four vector current density respectively.

6.17.5 Invariance of Maxwell's Equations under Lorentz Transformations

We have Maxwell's equations from classical electrodynamics point of view

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial_2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J} \quad \text{or} \quad \Box^2 \mathbf{A} = -\mu_0 \mathbf{J} \qquad (97)$$

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial_2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \quad \text{or} \quad \Box^2 \phi = -\frac{\rho}{\epsilon_0} \qquad (98)$$

Here $\mathbf{A} = (A_1, A_2, A_3)$ and $\mathbf{J} = (J_1, J_2, J_3)$

Classically, 'A' and 'ϕ', seems to be separate physical quantity, but from electrodynamics point of view, both are written as one physical quantity, called electromagnetic four-vector potential $(\mathbf{A}, \phi)$, and is written as, (A_1, A_2, A_3, ϕ). In context with electromagnetics, magnetic potential and electric potential are inter-connected with a well-known equation, called Lorentz condition, as

$$\nabla.\mathbf{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0$$

Actually, in electromagnetics $\frac{i\phi}{c} = \mathbf{A_4}$, is called fourth component of electromagnetic four-vector potential which is written as;

$A_\mu = (\mathbf{A}, \frac{i\phi}{c})$ here $\mu = 1,2,3,4$ with $\mathbf{A} = (A_1, A_2, A_3)$ and $\frac{i\phi}{c} = A_4$

In the light of above, equation (98) can be converted in terms of fourth component of electromagnetic potential $\mathbf{A_4}$ in the following way.

Multiplying equation (98) by $\frac{i}{c}$

$$\nabla^2 \frac{i\phi}{c} - \frac{1}{c^2}\frac{\partial_2}{\partial t^2}\left(\frac{i\phi}{c}\right) = -\frac{\rho}{\epsilon_0}\frac{i}{c}$$

$$\nabla^2 \frac{i\phi}{c} - \frac{1}{c^2}\frac{\partial_2}{\partial t^2}\left(\frac{i\phi}{c}\right) = -\mu_0 ic\rho \qquad \text{as } \frac{1}{c^2} = \mu_0\varepsilon_0 \Rightarrow \frac{1}{c\epsilon_0} = \mu_0 c$$

$$\nabla^2 \frac{i\phi}{c} - \frac{1}{c^2}\frac{\partial_2}{\partial t^2}\left(\frac{i\phi}{c}\right) = -\mu_0 J_4 \qquad \text{as } J_4 = ic\rho \qquad (99)$$

So, in the symmetry with equation (97), the equations (98) be written as.

$$\nabla^2 A_4 - \frac{1}{c^2}\frac{\partial_2}{\partial t^2} A_4 = -\mu_0 J_4$$

In general form, we can write;

$$\Box^2 A_\mu = -\mu_0 J_\mu \qquad \text{here } \mu = 1,2,3,4 \qquad (100)$$

Equation (100) is Maxwell's single equation of electromagnetism which is equivalent to two equations (97) and (98) and also equivalent to well-known four equations of electromagnetism. If A'_μ, J'_μ and $\Box'^2$ are the electromagnetic potential, electromagnetic current density and d'Alembert operator respectively, in four vector frame, w.r.t. relativistic moving frame of reference.

now for the Lorentz invariance, we have to prove

$$\Box'^2 A'_\mu = -\mu_0 J'_\mu \qquad (101)$$

We have to prove;

$$\Box^2 A_\mu = \Box'^2 A'_\mu \qquad (102)$$

And transformation is done using Lorentz transformation matrix as;

$$A'_\mu = \alpha_{\mu\nu} A_\mu$$

$$J'_\mu = \alpha_{\mu\nu} J_\mu$$

Where $\mathbf{A_\mu}$ and $\mathbf{J_\mu}$ are the electromagnetic potential and electromagnetic current density in four vector form in a frame at rest and $\alpha_{\mu\nu}$ is Lorentz transformation matrix, written as;

$$\alpha_{\mu\nu} = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix}, \quad \text{where } \beta = \frac{v}{c} \text{ and } \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

So, the four components of electromagnetic potential and current density can be found out as;

$$A'_\mu = \begin{bmatrix} A'_1 \\ A'_2 \\ A'_3 \\ A'_4 \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ \frac{i\phi}{c} \end{bmatrix}$$

$$J'_\mu = \begin{bmatrix} J'_1 \\ J'_2 \\ J'_3 \\ J'_4 \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & i\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\beta\gamma & 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} J_1 \\ J_2 \\ J_3 \\ ic\rho \end{bmatrix}$$

On simplifying the above matrix operations, we have,

$A'_1 = \gamma(A_1 - \frac{v}{c^2}\phi)$, $A'_2 = A_2$, $A'_3 = A_3$, $A'_4 = \gamma(\frac{i\phi}{c} - i\frac{v}{c}A_1) = \gamma(A_4 - i\frac{v}{c}A_1)$

$J'_1 = \gamma(J_1 - v\rho)$, (103)

$J'_2 = J_2$, and $J'_3 = J_3$, (104)

$J'_4 = \gamma(ic\rho - i\frac{v}{c}J_1) = \gamma(J_4 - i\frac{v}{c}J_1)$ (105)

To prove invariance, let us operate $\square'^2 A'_\mu$ for $\mu = 1,2,3,4$

(i) $\square'^2 A'_1 = \square^2[\gamma(A_1 - \frac{v}{c^2}\phi)]$, ($\square'^2 = \square^2$, as d'Alembert operator is invariant)

$$= \gamma[\square^2 A_1 - \frac{v}{c^2}\square^2\phi)],$$
$$= \gamma[\square^2 A_1 - \frac{v}{c^2}\square^2\phi)],$$
$$= \gamma[-\mu_0 J_1 + \frac{v}{c^2}\frac{\rho}{\epsilon_0})],$$ {putting the value of $\square^2 A_1$ and $\square^2\phi$, from (97) and (98)}
$$= \gamma(-\mu_0 J_1 + \mu_0 v\rho)$$
$$= -\mu_0\gamma(J_1 - v\rho)$$

or

$\square'^2 A'_1 = -\mu_0 J'_1$ as $J'_1 = \gamma(J_1 - v\rho)$, From (103)

(ii) $\square'^2 A'_2 = -\mu_0 J'_2$ and

(iii) $\square'^2 A'_3 = -\mu_0 J'_3$ As $J'_2 = J_2, J'_3 = J_3$, From (104)

(iv) $\square'^2 A'_4 = \square^2\gamma(A_4 - i\frac{v}{c}A_1)$

$$= \gamma(-\mu_0 J_4 + i\frac{v}{c}\mu_0 J_1)$$
$$= -\mu_0\gamma(J_4 - i\frac{v}{c}J_1),$$ or

$\square'^2 A'_4 = -\mu_0 J'_4$ as $J'_4 = \gamma(J_4 - i\frac{v}{c}J_1)$ from (105)

From equations (i), (ii), (iii) and (iv) can be generalized as;

$\square'^2 A'_\mu = -\mu_0 J'_\mu$ (106)

This proves the invariance of Maxwell's equations of electromagnetism.

*******Solved examples*******

Based on Galilean transformation

Ex.1 Use Galilean transformation to prove that the distance between two points (x_1, y_1, z_1) and (x_2, y_2, z_2) is invariant in two inertial frames.

Sol. Consider two frames S and S′, the latter moving with velocity v relative to former, such that $\mathbf{v} = \mathbf{i}v_x + \mathbf{j}v_y + \mathbf{k}v_z$.

Let the co-ordinates of two points in frame S be (x_1, y_1, z_1) and (x_2, y_2, z_2); while those in frame S′ be (x_1', y_1', z_1') and (x_2', y_2', z_2').
From Galilean transformation, we have

$$x_1' = x_1 - v_x t,\ y_1' = y_1 - v_y t,\ z_1' = z_1 - v_z t$$

and

$$x_2' = x_2 - v_x t,\ y_2' = y_2 - v_y t,\ z_2' = z_2 - v_z t$$

The distance between two points in frame S′

$$= \sqrt{\{(x_2' - x_1')^2 + (y_2' - y_1')^2 + (z_2' - z_1')^2\}}$$

Substituting the values of $x_1', x_2', y_1', y_2', z_1'$ and z_2', we get the distance between two points in S′.

$$= \sqrt{\left[\{(x_2 - v_x t) - (x_1 - v_x t)\}^2 - \{(y_2 - v_y t) - (y_1 - v_y t)\}^2 + \{(z_2 - v_z t) - (z_1 - v_z t)\}^2\right]}$$

$$= \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]}$$

= the distance between two points in system S.

Thus, we may say that the distance between any two points is invariant under Galilean transformations.

Ex.2 Water in a river moves east at 3 km/hour and a boat heads north at 4 km/hour with respect to water. Find the velocity of the boat with respect to ground and also find the direction.

Sol. Take east along +ve direction of x-axis and north along +ve y-axis. Consider the ground to be frame S while the water to be frame S′.

Then the velocity of the boat relative to frame S′, i.e., water $= u' = (4\text{km/hr}), \mathbf{j} = 4\mathbf{j}$, velocity of the system S′ relative to S i.e.,
velocity of the water relative to ground $= v = (3\text{km/hr}) \times \mathbf{i} = 3\mathbf{i}$.

According to transformation of velocity of the particle, we have

$$u' = u - v \quad \text{or } u = u' + v = 4\mathbf{j} + 3\mathbf{i}$$

Therefore, velocity of the boat relative to ground $= 3\mathbf{i} + 4\mathbf{j}$,

Magnitude of velocity $= |3\mathbf{i} + 4\mathbf{j}| = \sqrt{(3^2 + 4^2)} = 5\text{km/h}$

If α is the angle of the boat from east, we have

$$\cos\alpha = \frac{3}{|3i+4j|} = \frac{3}{5}$$

$$\alpha = \cos^{-1}\left(\frac{3}{5}\right)$$

Ex.3 A ball has velocity $(4\mathbf{i} - 5\mathbf{j} + 10\mathbf{k})$ m/s relative to a train moving with velocity $(3\mathbf{i} + 4\mathbf{j})$ m/s relative to an observer on the ground. Calculate the velocity of the ball relative to the ground.

Sol. Consider the train to be system S′ and the ground to be system S, then we have $\mathbf{v} = (3\mathbf{i} + 4\mathbf{j})\text{m/sec}$ and $\mathbf{u} = (4\mathbf{i} - 5\mathbf{j} + 10\mathbf{k})\text{m/sec}$

∴ The velocity of the ball relative to ground is given by

$$u = u' + v = (4\mathbf{i} - 5\mathbf{j} + 10\mathbf{k}) + (3\mathbf{i} + 4\mathbf{j})$$

$$= 7\mathbf{i} - \mathbf{j} + 10\mathbf{k}\ \text{m/sec}$$

Ex.4 Calculate the fringe-shift in Michelson-Morley experiment. Given that $l = 11$ metre, $v = 30\text{km/sec.}$ and $\lambda = 6 \times 10^{-5}$ cm.

Sol. The required fringe shift ΔN is given by $\Delta N = \frac{2lv^2}{c^2\lambda}$.

Here $\quad l = 11\text{m},\ \lambda = 6 \times 10^{-5}\ \text{cm} = 6 \times 10^{-7}\text{m}$

$\quad v = 30\ \text{km/s} = 3 \times 10^4\text{m/s},\ c = 3 \times 10^8\text{m/s}$

$\therefore \qquad \Delta N = \frac{2\times11\times(3\times10^4)^2}{(3\times10^8)^2\times6\times10^{-7}} = 0.37.$

Ex.5 Show by direct application of Lorentz transformations $x^2+y^2+z^2-c^2t^2$ is invariant.

Sol. We have only to prove by the help of Lorentz transformations,

$$x^2+y^2+z^2-c^2t^2 = x'^2+y'^2+z'^2-c^2t'^2$$

where (x,y,z,t) and (x',y',z',t') are the co-ordinates of the same event observed by two observers in systems S and S′ while S′ is moving with a velocity v relative to S.

Let us consider the expression

$$x'^2+y'^2+z'^2-c^2t'^2 = \frac{(x-vt)^2}{1-\frac{v^2}{c^2}}+y^2+z^2-c^2\frac{\left(t-\frac{vx}{c^2}\right)^2}{1-\frac{v^2}{c^2}}$$

[Putting values of x', y', z' and t' from Lorentz transformations]

$$x'^2+y'^2+z'^2-c^2t'^2 = \frac{c^2}{c^2-v^2}(x-vt)^2+y^2+z^2-\frac{c^2}{c^2-v^2}.c^2\left(t-\frac{vx}{c^2}\right)^2$$

$$= \frac{c^2}{c^2-v^2}\left[(x-vt)^2-c^2\left(t-\frac{vx}{c^2}\right)^2\right]+y^2+z^2$$

$$= \frac{c^2}{c^2-v^2}\left[x^2-2vxt+v^2t^2-c^2\left(t^2-\frac{2vxt}{c^2}+\frac{v^2x^2}{c^4}\right)\right]+y^2+z^2$$

$$= \frac{c^2}{c^2-v^2}\left[x^2-2vxt+v^2t^2-c^2t^2+2vxt-\frac{v^2x^2}{c^2}\right]+y^2+z^2$$

$$= \frac{c^2}{c^2-v^2}\left[\frac{c^2x^2+c^2v^2t^2-c^4t^2-v^2x^2}{c^2}\right]+y^2+z^2$$

$$= \frac{1}{c^2-v^2}[c^2(x^2-c^2t^2)-v^2(x^2-c^2t^2)]+y^2+z^2]$$

$$= \frac{1}{c^2-v^2}[(c^2-v^2)(x^2-c^2t^2)]+y^2+z^2$$

$$= x^2-c^2t^2+y^2+z^2 = x^2+y^2+z^2-c^2t^2$$

Thus, we have proved that, $x'^2+y'^2+z'^2-c^2t'^2 = x^2+y^2+z^2-c^2t^2$

The expression $x^2+y^2+z^2-c^2t^2$ is invariant under Lorentz transformations.

Based on length contraction

Ex.6 A rod has length 1 metre. When the rod is in a satellite moving with velocity 0.8c relative to laboratory, what is the length of the rod as determined by an observer (a) in the satellite and (b) in the laboratory?

Sol. (a) The observer in the satellite is at rest relative to the rod, therefore the length of the rod as measured by an observer in the satellite is 1 metre.

(b) The length of the rod in the laboratory is given by

$$l = l_0\sqrt{\left(1-\frac{v^2}{c^2}\right)}$$

Here $l_0 = 1\text{m}$, $v = 0.8c$, thus,

$$l = 1\sqrt{\left[1-\left(\frac{0.8c}{c}\right)^2\right]}$$

$$= 1\sqrt{(1-0.64)} = 1\sqrt{0.36} = 1\times0.6 = 0.6\text{ m}$$

Ex.7 The length of a rocket ship is 100 metre on the ground. When it is in flight its length observed on the ground is 99 metre, calculate its speed.

Sol. We have, $l = l_0\sqrt{\left(1-\frac{v^2}{c^2}\right)}$ or $\frac{l}{l_0} = \sqrt{\left(1-\frac{v^2}{c^2}\right)}$

Here $l_0 = 100$ metre, $l = 99$ metre

$$\frac{99}{100} = \sqrt{\left(1-\frac{v^2}{c^2}\right)} \Rightarrow \sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{99}{100}$$

or
$$v = \frac{\sqrt{99}}{100}c = \frac{\sqrt{99}}{100} \times 3 \times 10^8 \text{m/s}$$
$$= 4.23 \times 10^7 \text{m/s}$$

Ex.8 A rocket of proper length 600 meters is moving directly away from earth. A light pulse sent from the earth is reflected from mirrors at the rear (back part) and front of the rocket. If the first of these reflected pulses is received back 100 s after emission and the second one 16μs later, calculate

(a) the distance of the rocket from earth

(b) the velocity of the rocket and

(c) its apparent length

Sol. (a) The pulse sent from the earth reaches the rear of the rocket 50 s, after its emission, hence at this instant the distance of the rear of the rocket from earth $= 50 \times 3 \times 10^8$ m $= 1.5 \times 10^{10}$m

(b) According to Lorentz transformation in differential form

$$\delta t = \frac{\delta t' + \frac{v\,\delta x'}{c^2}}{\sqrt{1-\beta^2}}$$

If the events of reception of light pulse at the rear and front of the rocket be characterized by (x_1, t_1) and (x_2, t_2) in earth frame S and (x'_1, t'_1) and (x'_2, t'_2) in the rocket frame S′ moving relative to S along positive common x-axis, then given

$$\delta x' = x_2' - x_1' = 600 \text{ metre}$$
$$\delta t' = t_2' - t_1' = \frac{x_2' - x_1'}{c} = \frac{\delta x'}{c}$$

and
$$\delta t = t_2 - t_1 = \frac{16}{2}\mu s = 8\mu s = 8 \times 10^{-6} \text{ s}$$

∴
$$\delta t = \frac{\left(1+\frac{v}{c}\right)\delta x'}{c\sqrt{1-\frac{v^2}{c^2}}}$$

or
$$\sqrt{\left(\frac{1+\frac{v}{c}}{1-\frac{v}{c}}\right)} = \frac{c\,\delta t}{\delta x'}$$
$$= \frac{3 \times 10^8 \times 8 \times 10^{-6}}{600} = 4$$
$$v = 0.88\text{ c.}$$

(c)The apparent length of rocket is

$$l = 600\sqrt{\left(1-\frac{v^2}{c^2}\right)} = 60 \times \sqrt{[1-(0.88)^2]} = 280 \text{ metre.}$$

Ex.9 How fast would a rocket have to go relative to an observer for its length to be contracted to 99% of its length at rest?

Sol. According to length contraction

$$l = l_0\sqrt{\left(1-\frac{v^2}{c^2}\right)}$$

Here, $$l = \frac{99}{100} l_0$$

$$\Rightarrow \quad \frac{l}{l_0} = \frac{99}{100}$$

$$\therefore \quad \frac{99}{100} = \sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

$$\Rightarrow \quad \left(\frac{99}{100}\right)^2 = 1 - \frac{v^2}{c^2}$$

Or, $$\frac{v^2}{c^2} = 1 - \left(\frac{99}{100}\right)^2 = \frac{(100)^2 - (99)^2}{(100)^2} = \frac{199}{(100)^2}$$

$$v = \frac{\sqrt{199}}{100} c = 0.1416\ c.$$

Ex.10 A vector in system S′ is represented by 8**i** + 6**j**. How can the vector be represented in system S while S′ is moving with velocity 0.8 c**i** with respect to S; **i** and **j** being unit vectors along the respective directions.

Sol. S′ is moving relative to S with velocity 0.8c along x-axis; therefore, the length of the vector will only change along x-axis while its length along y-axis will remain the same.

We know, $$l = l_0 \sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

Here $$\frac{v}{c} = \frac{0.8c}{c} = 0.8.$$

$$\therefore \quad l_x = 8\sqrt{(1 - 0.64)} = 8\sqrt{(0.36)} = 8 \times 0.6 = 4.8$$

Therefore, in system S the vector may be represented by **4.8i + 6j.**

Ex.11 Calculate the percentage contraction of a rod moving with a velocity 0.8c in a direction inclined at 60° to its own length.

Sol. Let l_0 be length of the rod at rest.

If 0.8c is the velocity of the rod along the axis of x, then we have

$$\vec{l_0} = l_0 \cos 60° \, \mathbf{i} + l_0 \sin 60° \, \mathbf{j}$$

where **i** and **j** are unit vectors along the x and y-axis respectively.

As the contraction takes place in the direction of motion, the contraction will only take place along the axis of x; while the component of length along the axis of y will remain the same.

If l_x' is the component of the length of the rod, when in motion, along the axis of x, then we have

$$l_x' = l_x \sqrt{\left(1 - \frac{v^2}{c^2}\right)} = (l_0) \cos 60° \sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

$$= l_0 \cos 60° \sqrt{\left\{1 - \frac{(0.8c)^2}{c^2}\right\}} = \frac{l_0}{2} \times 0.6 \doteq 0.3 l_0$$

The component of the length of the moving rod along y-axis is given by

$$l_y' = l_0 \sin 60° = \frac{\sqrt{3}}{2} l_0.$$

$\therefore$ The length of the moving rod

$$= \sqrt{(l_x'^2 + l_y'^2)} = \sqrt{\left\{(0.3 l_0)^2 + \left(\frac{\sqrt{3} l_0}{2}\right)^2\right\}} = 0.911 l_0$$

$\therefore$ Decrease in length of the rod due to its motion

$$= l_0 - 0.91l_0 = 0.09l_0$$

Therefore, the percentage contraction $= 0.09\, l_0 \times \frac{100}{l_0} = 9\%$

Based on time dilation

Ex.12 The mean life of a meson is 2×10^{-6} s. Calculate the mean life of a meson moving with a velocity 0.8 c.

Sol. Given mean life $\tau = 2 \times 10^{-6}$ s and $v = 0.8\, c$

∴ Dilated mean life $\quad \tau' = \frac{\tau}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = \frac{2 \times 10^{-6}}{\sqrt{\left\{1-\left(\frac{0.8c}{c}\right)^2\right\}}}$

$$= \frac{2 \times 10^{-6}}{0.6} = 3.33 \times 10^{-8}\text{s}.$$

Ex.13 The half-life of a particular particle as measured in the laboratory comes out to be 4.0×10^{-8}s; when its speed is 0.8c and 3.0×10^{-8}s; when its speed is 0.6c. Explain this.

Sol. This can be explained on the basis of relativistic time dilation. The time interval in motion is given by

$$\Delta t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

where Δt is the proper time interval.

The proper half- life of the given particle is

$$\Delta t = \Delta t' \sqrt{\left(1-\frac{v^2}{c^2}\right)}$$

In first case given $\quad \Delta t' = 4.0 \times 10^{-8}$s and $v = 0.8\, c$

∴ $\Delta t = 4.0 \times 10^{-8} \sqrt{\left\{1-\left(\frac{0.8c}{c}\right)^2\right\}} = 4.0 \times 10^{-8} \times 0.6 = 2.4 \times 10^{-8}$s.

As proper half-life is independent of velocity, therefore half-life of the particle when speed is 0.6c must be given by:

$$\Delta t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = \frac{2.4 \times 10^{-8}}{\sqrt{\left\{1-\left(\frac{0.6c}{c}\right)^2\right\}}} = \frac{2.4 \times 10^{-8}}{0.8} = 3 \times 10^{-8}\text{s}$$

which is actual observation. Thus, the variation of half- life of the given particle is due to relativistic time dilation.

Ex.14 A beam of particles of half life 2×10^{-6}sec. travels in the laboratory with speed 0.96 times the speed of light. How much distance does the beam travel before the flux falls to ½ times the initial flux?

Sol. The observed half life $\Delta t'$ is given by $\Delta t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$

$$= \frac{2 \times 10^{-6}}{\sqrt{\left\{1-\left(\frac{0.96c}{c}\right)^2\right\}}} = \frac{2 \times 10^{-6}}{0.28} = 7.14 \times 10^{-6}\text{s}.$$

According to the definition of half- life in this observed time the flux falls to ½ times the initial values. Thus, in this time the distance traversed by the beam is

$$= v\Delta t' = 0.96 \times 3 \times 10^8 \times 7.14 \times 10^{-6} = 2.1 \times 10^3 \text{ metre}.$$

Ex.15 A certain particle called meson has a mean life time 2×10^{-6}s.
(a) What is the mean life time when the particle is travelling with speed of 2.99×10^8m/s.
(b) How far does it go during one mean life time.

Sol. (a) According to time dilation

$$\Delta t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

$$\Rightarrow \qquad \tau' = \frac{\tau}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}.$$

Here, proper mean life time $\tau = 2 \times 10^{-6}$s; $v = 2.99 \times 10^8$m/s

$\therefore$ Apparent mean life time of meson

$$\tau' = \frac{\tau}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

$$= \frac{2 \times 10^{-6}}{\sqrt{\left\{1-\left(\frac{2.99 \times 10^8}{3 \times 10^8}\right)^2\right\}}}$$

$$= \frac{2 \times 10^{-6}}{\sqrt{\left\{1-\left(\frac{2.99}{3}\right)^2\right\}}} = \frac{6 \times 10^{-6}}{\sqrt{(3)^2-(2.99)^2}}$$

$$= \frac{6 \times 10^{-6}}{\sqrt{5.99 \times 0.01}} = \frac{6 \times 10^{-6}}{0.245} = 2.45 \times 10^{-5}\text{s}.$$

(c) Distance traversed by beam in one mean life time

$$s = v\tau' = (2.99 \times 10^8\text{m/s}) \times (2.45 \times 10^{-5}\text{s})$$
$$= 7.325 \times 10^3\text{m} = 7.325 \text{ km}.$$

Ex.16 A wrist watch keeps correct time on the surface of earth. A pilot wears his watch and leaves the earth in a space ship which moves at a constant speed of 10^7m/s. How many seconds will the watch lose per day with respect to observer on earth?

Sol. According to time dilation

$$T' = \frac{T}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

If ΔT is the time lost by wrist watch per day, then proper time

$$T = T' - \Delta T$$

$$\therefore \qquad T' = \frac{T' - \Delta T}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

$$\therefore \qquad \frac{\Delta T}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = T'\left(\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} - 1\right) \Rightarrow \Delta T' = T'\left[\sqrt{\left(1-\frac{v^2}{c^2}\right)}\right]$$

Here $T' = 1$ day $= 24 \times 60 \times 60$s $= 86400$ s

Loss in time in seconds

$$\Delta T = 86400\left[1 - \sqrt{\left\{1-\left(\frac{10^7}{3 \times 10^8}\right)^2\right\}}\right]$$

$$= 86400\left[1-\sqrt{\left(1-\frac{1}{900}\right)}\right]$$

$$= 86400\left[1-\left(1-\frac{1}{900}\right)^{1/2}\right]$$

Using binomial theorem $\Delta T = 86400\left[1-\left(1-\frac{1}{2\times 900}\right)\right] = \frac{86400}{1800} = 48\text{ s}$

$$= 86400 \times 0.000667 = 58\text{ s.}$$

Ex.17 A particle with a mean proper life of 1 micro second (μs) moves through the laboratoty at 2.7×10^8 m/s.

(i) What will be its life time as measured by observer in the laboratory?

(ii) What will be the distance traversed by it before disintegrating?

(iii) Find the distance traversed without taking relativity into account.

Sol. (i) By the result of time dilation, we have

$$t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Given $\Delta t = 1\mu\text{ s} = 1 \times 10^{-6}\text{s}, v = 2.7 \times 10^8$ m/s

Therefore
$$\Delta t' = \frac{1\times 10^{-6}}{\left\{1-\left(\frac{2.7\times 10^8}{3\times 10^8}\right)^2\right\}}$$

$$= \frac{3\times 10^{-6}}{\sqrt{\{3^2-(2.7)^2\}}} = 2.3\times 10^{-6}\text{m} = 2.3\text{ μs}$$

(ii) Distance traversed by the particle

$$= v\,\Delta t' = 2.7\times 10^8 \times 2.3 \times 10^{-6}\text{m} = 620\text{m}$$

(iii) Distance traversed without relativistic effects

$$= v\,\Delta t = 2.7 \times 10^8 \times 1 \times 10^{-6} = 270\text{m}$$

Ex.18 The proper life of π^+-mesons is 2.5×10^{-8} s. If a beam of these mesons of velocity 0.8c is produced, calculate the distance, the beam can travel before the flux of the meson beam is reduced to $1/e^2$ times the initial flux.

Sol. According to time-dilation, we have

$$\Delta t' = \frac{\Delta t}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Given
$$\Delta t = 2.2 \times 10^{-8}\text{ s}, v = 0.8\text{ c.}$$

$$\Delta t' = \frac{2.5\times 10^{-8}}{\sqrt{\left\{1-\left(\frac{0.8c}{c}\right)^2\right\}}} = \frac{2.5\times 10^{-8}}{0.6} = 4.16\times 10^{-8}\text{s.}$$

If N_0 is the initial flux and N is the flux after time t, then we have

$$N = N_0 e^{-t/\tau}, \tau \text{ being mean life time}$$

Here
$$N = \frac{1}{e^2}N_0$$

∴
$$\frac{1}{e^2}N_0 = N_0 e^{-t/\tau}$$

or
$$t = 2\tau = 2\Delta t' \quad \text{as } (\tau = \Delta t' = 4.16\times 10^{-8}\text{s})$$

∴ The distance traversed by the beam before the flux of meson beam is reduced to $1/e^2$ times the initial flux.

$$= 2\Delta t' \times 0.8c = 8.32 \times 10^{-8} \times 0.8 \times 3 \times 10^8 = 19.96\text{ m.}$$

Ex.19 At what speed should a clock be moved so that it may appear to lose 1 minute in each hour?

Sol. If the clock is to lose 1 minute in one hour, then the clock must record 59 minutes for each hour recorded by clocks stationary with respect to the observer. If v is the required speed of the clock, then according to Einstein's apparent retardation of clocks, we must have

$$59 = 60\sqrt{\left(1-\frac{v^2}{c^2}\right)}$$

$$\frac{v}{c} = \sqrt{\left\{1-\left(\frac{59}{60}\right)^2\right\}} = 0.18$$

$$v = 0.18c = 5.4 \times 10^7 \text{m/s}.$$

Based on addition of velocities

Ex.20 An electron is moving with a speed of 0.8c in a direction opposite to that of a moving photon. Calculate the relative velocity of the electron and photon.

Sol. The speed of photon = c

The speed of electron, $v = 0.85c$

Let the photon and electron be moving along positive and negative directions of x-axis respectively. Let the electron moving with velocity $-0.85c$ be at rest in system S. Then we may assume that system S′ or laboratory is moving with velocity 0.85c relative to system S (electron). Then we may write $v = 0.85c, u_x' = c$.

$$\therefore \qquad u_x = \frac{u_x' + v}{1+\frac{u_x'v}{c^2}} = \frac{c+0.85c}{1+\frac{0.85c\,(c)}{c^2}} = \frac{(1+0.85)c}{(1+0.85)} = c.$$

Therefore, relative velocity of electron and photon = c.

Ex.21 Two particles come towards each other with speed 0.8c with respect to laboratory. What is their relative speed?

Sol. The problem consists of two velocities −0.8c and 0.8c observed from laboratory, therefore to find the required velocity we have to use the theorem of addition of velocities. For the solution of the problem consider a system S in which the particle having velocity −0.8c is at rest. Thus, we may say that system S′ i.e., laboratory is moving with velocity +0.8c relative to system S i.e., v = 0.8c. Now, we have to find the velocity of particle moving with velocity +0.8c relative to system S′ or laboratory in system S, we have from the theorem of addition of velocities.

$$u_x = \frac{u' + v}{1+\frac{u'v}{c^2}}.$$

Here $\qquad u' = 0.8c, v = 0.8c.$

Therefore, $$u_x = \frac{0.8c+0.8c}{1+\frac{(0.8c)(0.8c)}{c^2}} = \frac{0.8c+0.8c}{(1+0.64)} = \frac{1.6c}{1.64} = 0.9755c.$$

Thus, the relative speed of particles = 0.9755c.

Ex. 22 Two β particles A and B travel in opposite directions each with a velocity 0.9c. What is their relative velocity (i) as observed by A and (ii) as observed by a stationary observer?

Sol. In both cases (i) and (ii), $u' = 0.9c, v = 0.9c$

$$u_x = \frac{u' + v}{1+\frac{u'v}{c^2}} = \frac{0.9c+0.9c}{1+\frac{(0.9c)(0.9c)}{c^2}} = \frac{1.8}{(1+0.81)} = \frac{1.8}{1.81}c = 0.99c.$$

∴ Relative speed of particles in both cases (i) and (ii) is 0.99c.

Ex.23 A man leaves the earth in a rocket ship that makes a round trip to the nearest star which is 4 light years away at a speed of 0.8c. How much younger will he be on his return than his twin brother who preferred to stay behind.

Sol. According to length contraction, the apparent length of one-way journey.

$$l = l_0\sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

$$= 4 \times \sqrt{\left\{1 - \left(\frac{0.8c}{c}\right)^2\right\}}$$

$$= 4 \times 0.6\, l_y = 2.4 l_y$$

Total apparent length of round trip L = 2l = 4.8 light years.

∴ Total time interval of man on trip

$$t = \frac{\text{distance}}{\text{speed}}$$

$$= \frac{4.8 \text{ light years}}{0.8c} = 6 \text{ years}$$

According to time dilation,

$$\text{dilated time interval, } \Delta T' = \frac{T}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$$

$$T' = \frac{6}{\sqrt{\left(1 - \left(\frac{0.8c}{c}\right)^2\right)}} \text{years} = \frac{6}{0.6} = 10 \text{ years.}$$

Apparent age difference between two brothers

$$\Delta T = T' - T$$

$$10 - 6 = 4 \text{ years.}$$

Hence the man on trip will appear to be 4 years younger than his twin brother (who was at rest on earth).

Ex. 24 A scientist observes that a certain atom A moving relative to him with velocity $2 \times 10^8 m/s$ emits a particle B which moves with velocity $2.8 \times 10^8 m/s$ with respect to atom. Calculate the velocity of the emitted particle relative to scientist.

Sol. According to velocity addition theorem, we have

$$u_x = \frac{u' + v}{1 + \frac{u'v}{c^2}}$$

Let us consider the particle A to be in system S' and scientist in system S.

Then u' = velocity of the particle B relative to S'= $2.8 \times 10^8 m/s$.

v = velocity of system S' relative to S =$2 \times 10^8 m/sec$.

u' = velocity of the particle B relative to system S = ?

$$u' = \frac{2.8 \times 10^8 + 2 \times 10^8}{1 + \frac{(2.8 \times 10^8) \times (2 \times 10^8)}{(3 \times 10^8)^2}} = \frac{4.8 \times 10^8}{1 + \frac{5.6}{9}} = \frac{4.8 \times 10^8}{1.622}$$

$$= 2.95 \times 10^8 m/s.$$

Ex.25 A rocket is chasing enemy's space-ship. An observer on the earth observes the speed of rocket to be $2.5 \times 10^8 m/s$ and that of space-ship $2 \times 10^8 m/sec$. Calculate (i) the velocity of enemy's ship as seen by rocket

Sol. According to velocity addition theorem $u' = \frac{u'+v}{1+\frac{u'v}{c^2}}$

Consider the rocket to be system S and the earth system S′. As rocket is moving relative to earth with speed $2.5 \times 10^8 m/s$, it may be assumed that the earth (S′) is moving relative to rocket (S) with velocity $-2.5 \times 10^8 m/s$, i.e., $v = -2.5 \times 10^{10} m/s$ and u' = velocity of enemy's ship relative to earth $S' = 2 \times 10^8 m/s$.

∴ The velocity of enemy's ship as seen by rocket,

$$u_x = \frac{u'+v}{1+\frac{u'v}{c^2}} = \frac{2\times 10^8 - 2.5\times 10^8}{1+\frac{(2\times 10^8)\times(-2.5\times 10^8)}{(3\times 10^8)^2}} = \frac{-0.5\times 10^8}{1-\frac{5}{9}} = -\frac{0.5\times 10^8}{0.44}$$

$$= -1.14 \times 10^8 m/s.$$

i.e., the observer on the rocket will observe that enemy's ship is approaching him with velocity $1.14 \times 10^8 m/s$.

Based on relativistic mass

Ex.26 Find the velocity at which the mass of a particle is double its rest mass.

Sol. The mass of a particle (rest mass m_0) moving with velocity v is given by

$$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Given $\quad m = 2m_0$

∴ $\quad 2m_0 = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} \quad$ or $\quad 2 = \frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$

or $\quad 1-\frac{v^2}{c^2} = \frac{1}{4} \quad$ or $\quad \frac{v^2}{c^2} = \frac{3}{4}$

∴ $\quad v = \frac{\sqrt{3}}{2}c = \frac{1.732}{2} \times 3 \times 10^8 m/s = 2.598 \times 10^8 m/s.$

Ex.27 At what speed will the mass of a body is 2.25 times its rest mass?

Sol. According to variation of mass with velocity

$$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} \Rightarrow \frac{m}{m_0} = \frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Given $\quad m = 2.25\, m_0 \Rightarrow \frac{m}{m_0} = 2.25$

∴ $\quad 2.25 = \frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$

∴ $\quad \sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{1}{2.25}$

$$1-\frac{v^2}{c^2} = \left(\frac{1}{2.25}\right)^2 \Rightarrow \frac{v^2}{c^2} = 1-\frac{16}{81}$$

or $\quad \frac{v^2}{c^2} = \frac{65}{81} \quad \Rightarrow \sqrt{\frac{65}{81}}\, c = 0.895c$

or $\quad v = 0.895 \times 3 \times 10^8 = 2.68 \times 10^8 m/s.$

Ex.28 How fast must an electron move in order that its rest mass becomes equal to the rest mass of a proton?

Sol. According to formula of mass variation with velocity

$$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Here Mass of m = mass of proton = 1.67×10^{-27}kg

Mass of electron = 9.1×10^{-31} kg

$$\therefore \quad 1.67 \times 10^{-27} = \frac{9.1 \times 10^{-31}}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

$$\Rightarrow \quad \sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{9.1 \times 10^{-31}}{1.67 \times 10^{-27}}$$

$$\text{or} \quad \sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{9.1}{1.67} \times 10^{-4}$$

$$1-\frac{v^2}{c^2} = \left(\frac{9.1}{1.67} \times 10^{-4}\right)^2$$

$$\frac{v^2}{c^2} = 1-\left(\frac{9.1}{1.67} \times 10^{-4}\right)^2$$

$$v \approx c$$

That is the electron must move with speed very slightly less than the speed of light.

Ex.29 A man weighs 50 kg on the earth. When he is in a rocket ship in flight, his mass is 50.5 kg as measured by an observer on earth. What is the speed of rocket?

Sol. From mass variation relation,

$$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Given mass in motion, m = 50.5 Kg, Rest mass, m_0 = 50 Kg

$$\therefore \quad 50.5 = \frac{50}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} \quad \Rightarrow \quad \sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{50}{50.5}$$

$$\Rightarrow \quad \frac{v^2}{c^2} = 1-\left(\frac{50}{50.5}\right)^2$$

$$\Rightarrow \quad v = \sqrt{0.0197}\, c$$

$$= 0.138 \times 3 \times 10^8$$

$$= 4.13 \times 10^7 \text{m/s.}$$

Ex.30 The rest mass of an electron is 9.1×10^{-31} kg. What will be its mass if it were moving with 4/5th the speed of light?

Sol. The mass of the electron if it were moving with speed v is given by $m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$,

where m_0 is the rest mass of the electron.

Here $v = \frac{4}{5}c = 0.8c$ and $m_0 = 9.1 \times 10^{-31}$ Kg

$$\therefore \quad m = \frac{9.1 \times 10^{-31}}{\sqrt{\left\{1-\left(\frac{0.8c}{c}\right)^2\right\}}} = \frac{9.1 \times 10^{-31}}{\sqrt{1-0.64}} = \frac{9.1 \times 10^{-31}}{\sqrt{0.36}}$$

$$= \frac{9.1 \times 10^{-31}}{0.6} = 15.16 \times 10^{-31} \text{ Kg}$$

Based on Einstein mass energy equivalence

Ex.31 Calculate the rest energy of electron and proton in MeV. Given $m_e = 9.11 \times 10^{-31}$ kg and $m_p = 1.673 \times 10^{-27}$**kg.**

Sol. Einstein mass-energy equivalence relation is $E = mc^2$.

For electron, $m_e = 9.11 \times 10^{-31}$ Kg

$$E_e = m_e c^2 = 9.11 \times 10^{-31} \times (3 \times 10^8)^2 \text{Joule}$$

$$= \frac{(9.11 \times 10^{-31}) \times ((3 \times 10^8)^2)}{1.6 \times 10^{-13}} \text{ MeV} = 0.512 \text{ MeV}$$

For proton, $m_p = 1.673 \times 10^{-27}$ Kg

$$E_p = m_p . c^2 = 1.673 \times 10^{-27} \times (3 \times 10^8)^2 \text{Joule}$$

$$= \frac{1.673 \times 10^{-27} \times (3 \times 10^8)^2}{1.6 \times 10^{-13}} \text{ MeV} = 941 \text{MeV}$$

Ex.32 Find the mass and speed of 2 MeV electrons. Rest mass of electron $= 9.11 \times 10^{-31}$kg.

Sol. Given $E = 2 \text{ MeV} = 2 \times 10^6 \times 1.6 \times 10^{-19} = 3.2 \times 10^{-13}$ J

From Einstein's mass – energy equivalence relation $E = mc^2$, we have

Mass of electron, $m = \frac{E}{c^2}$

$$= \frac{3.2 \times 10^{-13}}{(3 \times 10^8)^2} = 3.55 \times 10^{-30} \text{Kg}$$

From mass-velocity relation $m = \frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$

$$\frac{m_0}{m} = \sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

$$\frac{9.11 \times 10^{-31}}{3.55 \times 10^{-30}} = \sqrt{\left(1 - \frac{v^2}{c^2}\right)}$$

$\Rightarrow$ $\sqrt{\left(1 - \frac{v^2}{c^2}\right)} = 0.256$

$\Rightarrow$ $\frac{v^2}{c^2} = 1 - (0.256)^2 = 0.9343$

$\Rightarrow$ $v = 0.966$ c.

Ex.33 How does a proton gain in mass when accelerated to a kinetic energy of 500 MeV?

Sol. Gain in kinetic energy of proton

$$\Delta E = 500 \text{ MeV}$$

$$= 500 \times 10^6 \times 1.6 \times 10^{-19} \text{J}$$

$$= 8.0 \times 10^{-11} \text{ J}$$

According to Equivalence of mass-energy relation

$$\Delta E = \Delta m . c^2$$

∴ Gain in mass, $\Delta m = \frac{\Delta E}{c^2}$

$$= \frac{8.0 \times 10^{-11}}{(3 \times 10^8)^2} = 8.89 \times 10^{-28} \text{ Kg}$$

Ex.34 A particle of mass m_0 moves with speed $\frac{c}{\sqrt{2}}$. Calculate the mass, momentum, total energy and kinetic energy of the particle.

Sol. Mass of particle $$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Here $$v = \frac{c}{\sqrt{2}}$$

∴ Mass of particle, $$m = \frac{m_0}{\left(1-\frac{(c/\sqrt{2})^2}{c^2}\right)} = \frac{m_0}{\sqrt{1-\frac{1}{2}}} = \sqrt{2}\, m_0$$

Momentum, $$p = mv = (\sqrt{2}\, m_0)\frac{c}{\sqrt{2}} m_0 c$$

Total energy, $$E = mc^2 = (\sqrt{2}\, m_0)c^2 = \sqrt{2}\, m_0 c^2$$

Kinetic energy, $$T = \text{Total energy} - \text{rest energy}$$
$$= mc^2 - m_0c^2$$
$$T = \sqrt{2}\, m_0c^2 - m_0c^2$$
$$= 1.41 m_0c^2 - m_0c^2$$
$$= 0.41\, m_0c^2.$$

Ex.35 If the kinetic energy of a body is twice its rest mass energy, find its velocity.

Sol. Rest mass energy = m_0c^2

Relativistic kinetic energy $$T = (m - m_0)c^2$$

Given $$T = 2m_0c^2$$

∴ $$(m - m_0)c^2 = 2m_0c^2 \Rightarrow m = 3m_0 \quad (1)$$

As $$m = \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

Equation (1) gives $$= \frac{m_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = 3m_0$$

or $$\sqrt{\left(1-\frac{v^2}{c^2}\right)} = \frac{1}{3}$$

$$1 - \frac{v^2}{c^2} = \frac{1}{9}$$

or $$\frac{v^2}{c^2} = \frac{8}{9}$$

$$v = \frac{2\sqrt{2}}{3}c.$$

Ex.36 Calculate the fractional change in the mass of the hydrogen atom when it is ionized from the following data:

The binding energy of hydrogen atom= 13.6 eV

The rest mass of hydrogen atom, $m_0 = 1.00797$ a.m.u

$$1 \text{ amu} = 1.66 \times 10^{-27} \text{ Kg}$$

Sol. The change in rest mass $\Delta m_0 = \frac{\Delta E}{c^2}$.

$$= \frac{13.6 \text{ eV}}{(3\times 10^8)^2} = \frac{13.6 \times 1.6 \times 10^{-19}\text{joule}}{(3\times 10^8)^2}$$

$$= \frac{13.6 \times 1.6\times 10^{-19}\text{joule}}{(3\times 10^8)^2}\text{Kg} = 2.42 \times 10^{-35}\text{Kg}$$

$$\doteq \frac{2.42 \times 10^{-35}}{1.66 \times 10^{-27}} \text{ a.m.u} \doteq 1.46 \times 10^{-8} \text{ a.m.u}$$

So, fractional change in the mass of the hydrogen atom,

$$\therefore \quad \frac{\Delta m_0}{m_0} = \frac{1.46 \times 10^{-8}}{1.00797} = 1.45 \times 10^{-8}$$

Ex.37 Find the velocity that an electron must be given so that its momentum is 10 times its rest mass times the speed of light. What is the energy at this speed?

Sol. The momentum of an electron of rest mass m_0 moving with velocity v is given by

$$p = \frac{m_0 v}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$$

Given $\quad \frac{m_0 v}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} = 10\, m_0 c \quad \text{or} \quad \frac{v}{c} = 10\sqrt{\left(1 - \frac{v^2}{c^2}\right)}$

Squaring we get $\quad \frac{v^2}{c^2} = 100\left(1 - \frac{v^2}{c^2}\right) = 100 - 100\frac{v^2}{c^2}$

or $\quad (100 + 1)\frac{v^2}{c^2} = 100$

or $\quad \frac{v^2}{c^2} = \frac{100}{101} \quad (1)$

or $\quad \frac{v}{c} = \sqrt{\frac{100}{101}} = 0.995$

$\therefore$ $v = 0.995c = 0.995 \times 3 \times 10^8 m/s = 2.985 \times 10^8$ m/s

The mass of the electron at this speed is given by $m = \frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$

where m_0 = rest mass of electron $= 9 \times 10^{-31}$kg.

$$m = \frac{9 \times 10^{-31}}{\sqrt{\left(1 - \frac{100}{101}\right)}} \qquad \left(\text{since} \frac{v^2}{c^2} = \frac{100}{101} \text{from}(1)\right)$$

$$= \frac{9 \times 10^{-31}}{\sqrt{\left(\frac{1}{101}\right)}} = 9 \times 10^{-31}\sqrt{(101)}$$

$$= 9 \times 10^{-31} \times 10.04 = 90.36 \times 10^{-31} \text{ Kg} = 9.036 \times 10^{-30} \text{ Kg}$$

$\therefore$ Energy of the electron at speed 0.995c is

$$E = mc^2 = (9.036 \times 10^{-30} kg)(3 \times 10^8 m/s)^2$$
$$= 8.13 \times 10^{-13} \text{ J}$$

Ex.38 The mass of a moving electron is 11 times its rest mass. Calculate its kinetic energy and momentum.

Sol. Rest mass of electron $m_0 = 9 \times 10^{-31}$ Kg

Let 'm' be the mass of electron when it moves with speed v, then

$$m = \frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} \quad (1)$$

Given $\quad m = 11m_0$

$\therefore \quad 11m_0 = \frac{m_0}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$

$$\Rightarrow \qquad \sqrt{\left(1-\frac{v^2}{c^2}\right)}=\frac{1}{11}$$

or
$$1-\frac{v^2}{c^2}=\frac{1}{121}$$

$$\Rightarrow \qquad \frac{v^2}{c^2}=1-\frac{1}{121} \text{ or } \frac{v}{c}=\sqrt{\frac{120}{121}}$$

$$\Rightarrow \qquad v=\sqrt{\frac{120}{121}}c=\sqrt{\frac{120}{121}}\times 3\times 10^8 m/s=2.975\times 10^8\ m/s$$

Kinetic energy of electron, $K\,E=(m-m_0)c^2$

$$=(11m_0-m_0)c^2=10m_0c^2$$
$$=10\times 9.1\times 10^{-31}\times(3\times 10^8)^2\ J$$
$$=8.19\times 10^{-13}\ J$$

Momentum of electron $p=mv=11m_0v$

$$=11\times 9.1\times 10^{-31}\times 2.975\times 10^8$$
$$=2.978\times 10^{-21}\ Kg\ m/s$$

Ex.39 An electron and a positron practically at rest come together and annihilate each other, producing two photons of equal energy. Find the energy and equivalent mass of each photon.

Sol. The rest mass of electron $=9\times 10^{-31}$ Kg = rest mass of positron

∴ The energy of each photon = The energy equivalent to each particle $=m_0c^2$

$$=9\times 10^{-31}(3\times 10^8)^2\ J$$
$$=81\times 10^{-15}\ J=\frac{81\times 10^{-15}}{1.6\times 10^{-19}}eV=5\times 10^5 eV$$

The equivalent mass of each photon $=9\times 10^{-31}$ Kg

Ex.40 Calculate the velocity of electrons accelerated by a potential of 1 million-volts.

Sol. We know, that kinetic energy, $T=(m-m_0)c^2=\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}m_0c^2$

Given $T=1\ MeV=10\ eV=10^6\times 1.6\times 10^{-19} J, m_0=9\times 10^{-31} Kg$

$$\therefore \qquad 10^6\times 1.6\times 10^{-19}=\left[\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}\right]\times 9\times 10^{-31}\times(3\times 10^8)^2$$

Solving we get
$$v=\frac{2\sqrt{2}}{3}c=\frac{2\sqrt{2}}{3}\times 3\times 10^8= m/s=2.82\times 10^8\ m/s$$

Ex.41 A hundred μ mesons, each of rest mass 206 electron masses and energy 4.75 BeV are produced at an altitude of 30 km. If the mean life of μ-mesons at rest is 2.2×10^{-6} second, calculate their number expected to reach the sea-level (a) allowing for time dilation and (b) neglecting time dilation. Take the electrons to travel vertically downwards without loss of energy. Given the electron rest mass = 0.5 MeV.

What conclusion can you draw from your result?

Sol. The rest mass of a μ-meson is

$$m_0c^2=206\times 0.5\ MeV=103\ MeV=0.103\ BeV$$

The kinetic energy of a μ-meson is $T=(m-m_0)c^2=4.75$ BeV

$$\left[\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}-1\right]m_0c^2=4.75\ BeV$$

or $$\left[\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}-1\right] \times 0.103 \text{ BeV} = 4.75 \text{ BeV}$$

or $$\left[\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}\right] = \frac{4.75}{0.103}+1 = 47.12$$

or $$\frac{1}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = \frac{1}{47.12}$$

or $$\frac{v}{c} = \sqrt{\left\{1-\left(\frac{1}{47.12}\right)^2\right\}} \approx 1$$

$\therefore$ $$v \approx c$$

(a) Allowing time dilation, the average life of moving μ-mesons is

$$\tau = \frac{\tau_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}} = 47.12 \times 2.2 \times 10^{-6} = 1.04 \times 10^{-4}\text{s}.$$

If N is the number of μ − mesons reaching the sea level undecayed, we have

$$N = N_0 e^{-t/\tau}$$

Here $$t = \frac{30 \times 10^3}{3 \times 10^8} = 10^{-4}\text{s}.$$

$\therefore$ $$N = 100\, e^{\frac{-10^4}{1.04} \times 10^{-4}} = 100e^{-1/1.04} \approx 38.$$

(b) Neglecting time dilation, the mean life of moving μ-mesons is to be taken as 2.2×10^{-6} second. Thus the number of μ-mesons reaching the sea-level undecayed is

$$N = N_0 e^{-\frac{t}{\tau_0}} = 100e^{\frac{-10^{-4}}{2.2} \times 10^{-6}} \approx 1.7 \times 10^{-18}.$$

Conclusion: It is obvious from the results availability of cosmic rays μ-mesons at sea level can be explained only on the basis of relativistic time dilation.

*****Review Questions and Problems*****

Based on frames of reference, Galilean transformations

1. What do you mean by a frame of reference?
2. Define an inertial frame and explain the principle of the special theory of relativity. Give an instance to show that this principle leads to revision for our ordinary concepts of space and time.
3. What do you mean by inertial and non- inertial frames of reference?
4. Give an example of non-inertial frame of reference.
5. Is earth an inertial frame? If not why?
6. Write Galilean transformations for space and time.
7. What are inertial frames? Discuss Galilean transformation for position, velocity and acceleration.
8. Prove that the Newton's laws of motion are invariant under Galilean transformations?

9. What are Galilean transformations? Obtain the transformation equations for velocity and acceleration in frame S' moving with velocity v relative to inertial frames

Based on Michelson Morley experiment

10. With what aim in mind, the Michelson-Morley experiment was performed? State its conclusions.
11. Describe the Michelson—Morley experiment and explain the physical significance of the negative result.
12. Give the main conclusion of the Michelson-Morley experiment.
13. Describe the Michelson-Morley experiment. Why is Michelson and Morley experiment considered to be in contradiction to ether hypothesis?
14. In an experiment, the length of the arm of the interferometer was 11 metres, the wavelength of light 5.5 x 10^{-7} metre and earth's velocity 30 km/s., calculate the amount of fringe shift. [Ans. 0.4 fringes]

Based on Length contraction

15. What is proper length? Explain how Lorentz transformations account for the phenomenon of length contraction.
16. Explain Lorentz-Fitzgerald contraction idea. How was this idea used to account for the negative result of the Michelson-Morley experiment.
17. What do you mean by length contraction at relativistic speed? Deduce the necessary expression for it? Calculate the percentage contraction of a rod moving with a velocity 0.8 times the velocity of light in a direction inclined at 45° to its own length. [Ans: 17.5 %]
18. A beam of particles of half-life 2.0 x 10^{4} s travels in the laboratory with speed 0.96c. How much distance does the beam travel before the flux of the beam falls to half the initial flux?
19. The length of a rod is 100m. If the length of this rod is measured by the observer moving parallel to its length is 51m, find the speed of the observer.
20. A rod has a length of one metre. It is placed in a space ship moving with a velocity 0.4c relative to the earth. Calculate its length as measured by an observer (i) in spaceship (ii) on earth. [Ans. (i) 1 m (ii) 0.92 m]
21. What will be the length of a metre rod appear to be for a person travelling parallel to the length of the rod at a speed of 0.8c relative to the rod. [Ans. 0.6 m]
22. A certain particle has a life time of 10^{-7} sec, when measured at rest. How far does it go before decaying if its speed is 0.99c when it is created? [Ans. 208.97 m]

Based on Time dilation

23. What do you mean by time dilation?
24. Write down Lorentz transformation equations and hence explain Lorentz Fitzgerald contraction and time dilation.
25. Explain the terms proper and non-proper time intervals as used in special relativity. Describe an experiment to verify time dilation.
26. Give experimental verification of the phenomenon of time-dilation
27. Derive the Lorentz transformation equations. Using them prove that "moving clocks appear to go slow".

28. (a) What must be one's speed, relative to a frame S, in order that one's clock will lose 1 second per day as observed from S? (b) What if it is to lose 1 min per day? [Ans. (a) 1.4×10^6 m/s, (b) 1.1×10^7 m/s]
29. The half-life of an elementary particle as measured in the laboratory is 4.0×10^{-8} s when its speed is 0.8c. Calculate (a) its proper half-life (b) its half-life at speed 0.6c. [Ans. (a) 2.4×10^{-8} s, (b) 3.0×10^{-8} s]
30. A beam of unstable elementary particles travels at a speed of 0.9c. At this speed, the mean life as measured in the laboratory frame is 5×10^4 s. Calculate the proper mean life of the particles. [Ans. 2.18×10^{-6} s]
31. A certain process requires 10^{-6} s to occur in an atom at rest in laboratory. How much time will it take to an observer in laboratory if the atom is moving with velocity 2×10^8 m/s. [Ans. 1.23×10^{-6} s]

Based on Postulates of special theory of relativity

32. What are postulates of special theory of relativity?
33. State the fundamental postulates of special theory of relativity and deduce from them the Lorentz transformations and show how these are superior to Galilean transformations.
34. Show that $x^2 + y^2 + z^2 - c^2t^2$ is invariant under Lorentz transformation.
35. Apply Lorentz transformation to derive expressions for length contraction and time dilation.
36. Using the Lorentz transformation, obtain the relativistic law of addition of velocities.
37. Derive Lorentz transformation equations for space and time coordinates and show that these equations become the Galilean equations at very low speeds.
38. Apply Lorentz transformation to derive expression for length contraction and time dilation.
39. Deduce the equations for transformation of velocity using Lorentz transformations. Show that velocity of light is invariant under Lorentz transformations.

Based on Law of addition of velocities

40. Using Lorentz transformations, obtain the law of addition of velocities. From this, show that a photon moving with velocity 'c' in one system S will appear to move with the same velocity 'c' in another system S' which is moving with a velocity relative to S.
41. Deduce the relativistic velocity addition theorem. Show that it is consistent with Einstein's second postulate of special theory of relativity.
42. Deduce the relativistic velocity addition theorem. Hence show that no signal can travel faster than light.
43. Show that no two velocities can be added to a value of more than the velocity of light (c) of light.
44. Spacecraft A is moving at 0.9c relative to the earth. Another spacecraft B passes A at a relative speed of 0.5c in the same direction. What is the speed of B relative to the earth? [Ans. 0.97c]

45. Two spaceships A and B are moving in opposite directions. An observer on the earth measures the speed of A to be 0.75c and the speed of B to be 0.85e. Find the velocity of B relative to A. [Ans. 0.9771c]
46. A rocket travelling away from the earth with a speed of 0.5 c fires off a second rocket at a speed of 0.6 c with respect to the first one. Calculate the speed of the second rocket with respect to the earth.
47. Two spaceships leave the earth in opposite directions, each with a speed of 0.5c with respect to the earth. What is the velocity of spaceship 1 relative to spaceship?

Based on Mass energy equivalence

48. What do you mean by mass-energy equivalence relation?
49. What is maximum possible velocity of a material particle?
50. What is the rest mass of the photon?
51. Explain and establish mass energy equivalence $E = mc^2$.
52. What are consequences if momentum and energy are invariant under Lorentz transformations? Prove the relation $E^2 - p^2c^2 = m^2c^4$, where symbols have their usual meaning.
53. Write a note on Einstein mass energy relation. What is the principle of mass and energy equivalence? Explain it by giving some examples.
54. What are the two important postulates of special theory of relativity? Deduce the mass-energy relation.
55. Can a particle move through free space at a speed greater than the speed of light. Justify your answer.
56. Derive an expression for the variation of mass with velocity.
57. Show that a particle having zero rest mass is always moving at the speed of light.
58. What should be the speed of electron so that its relativistic mass is twice its rest mass')
59. State the formula for the variation of mass with velocity. Compute to three significant places the speed at which the mass of a body becomes 8 times of that at rest ($c = 3.00 \times 10^8$ m/s). [Ans. 2.98×10^8 m/s]
60. Compute the speed at which mass of an electron becomes 4 times its rest mass. What is the physical significance of this increase of mass? [Ans. 2.94×10^8 m/s]
61. Calculate the velocity of a particle at which its mass will become 8 times its rest mass. [Ans. 0.992 c]

****