MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2024-25

Applied Physics Notes

Sl. No.	Faculty Name	Designation
1	Dr. Shivalinge Gowda	Professor & Head
2	Ms. Sushasini N K	Assistant Professor

MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2022-23

Module – 1 : Laser & Optical Fibers

Notes

Syllabus:

Laser and Optical Fibers: **8 Hrs**

LASER: Characteristic properties of a LASER beam, Interaction of Radiation with Matter, Einstein's A and B Coefficients and

Expression for Energy Density (Derivation), Laser Action, Population Inversion, Metastable State, Requisites of a laser system,

Semiconductor Diode Laser, Applications: Bar code scanner, Laser Printer, Laser Cooling (Qualitative), Numerical Problems.

Optical Fiber: Principle and Structure, Propagation of Light, Acceptance angle and Numerical Aperture (NA), Derivation of Expression for NA, Modes of Propagation, RI Profile, Classification of Optical Fibers, Attenuation and Fiber Losses,

Applications: Fiber Optic networking, Fiber Optic Communication. Numerical Problems

Pre requisite: Properties of light

Self - learning: Total Internal Reflection

LASER

Introduction:

LASER is the acronym (short form) for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. The important Characteristic properties of LASER beam are

1. It is highly monochromatic.
2. It is highly coherent.
3. It is highly directional.
4. It is a high intensity beam of light.

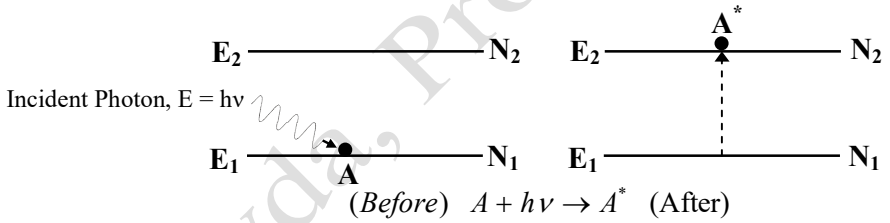
In order to understand the mechanism involved in the production of laser beam, one has to know the process taking place in an atomic system such as absorption and Emission of radiation.

Interaction of radiation with matter:

The interaction of radiation with matter occurs through the following three processes, namely.

1. Induced absorption
2. Spontaneous emission and
3. Stimulated emission

1. Induced Absorption:



When a suitable energy of a photon is incident on an atom, the photon is absorbed it. In this process the incident photon excites an atom from ground state to higher energy sate and hence it is known as Induced Absorption.

Consider an atom in a lower energy states E_1, it will excite to higher energy states E_2 by absorbing the incident photon of energy $E = h\nu = E_2 - E_1$. where h is the Planck's constant and ν is the frequency of the incident photon. The induced absorption can be represented as

$$A + h\nu \rightarrow A^*$$

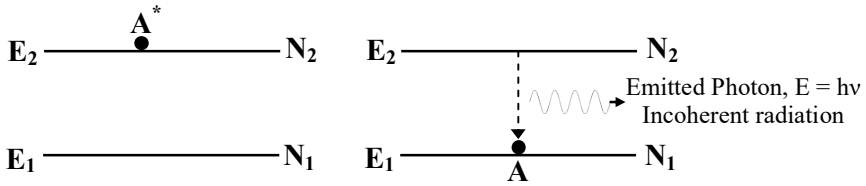
Where A is the atom in the ground sate E_1 and A* is the excited atom in the higher energy sate E_2.

Let N_1 and N_2 be the number of atoms in the energy levels E_1 and E_2, and U_ν be the energy density of the incident radiation. Thus the probability of transition of atoms from E_1 to E_2 is depending on N_1 and U_ν .

Therefore, the rate of induced absorption is $\propto N_1U_\nu = B_{12}N_1U_\nu$

Where B_{12} is a proportionality constant known as the Einstein's coefficient for induced absorption.

2. Spontaneous emission:



(*Before*) $A^* \rightarrow A + h\nu$ (After)

Spontaneous emission is the process of emission of photon, when an atom transits from higher energy level to lower energy level without the influence of any external energy.

An atom in the higher energy state E_2 makes a transition to lower energy state E_1 without the action of any external agency. the photon of energy $E_2 - E_1 = h\nu$ is emitted. In this process the emitted photons need not travel in the same direction. Thus the emitted light beam is not directional. The spontaneous emission can be represented as

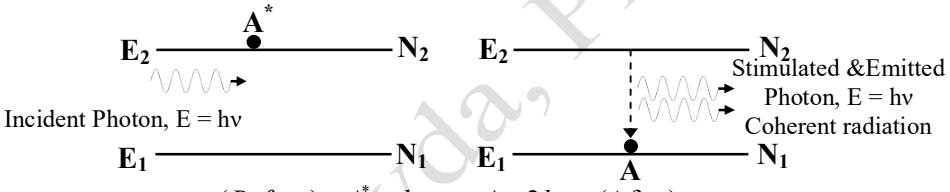
$$A^* \rightarrow A + h\nu$$

In this process, the probability of transition of atoms from E_2 to E_1 is depending on N_2 only.

Therefore, the rate of spontaneous emission is $\propto N_2 = A_{21}N_2$

Where A_{21} is a proportionality constant known as the Einstein's coefficient for spontaneous emission.

3. Stimulated emission:



(*Before*) $A^* + h\nu \rightarrow A + 2h\nu$ (After)

When a photon of suitable energy interacts with an atom in the higher energy state without loose energy then the atom is stimulated (Forced) to make transition from higher energy state E_2 to the lower energy state E_1 with the emission of a photon of energy $E_2 - E_1 = h\nu$. Both the incident photon and the emitted photons are coherent or in phase and travel in the same direction. This process is known as stimulated emission. The stimulated or induced emission can be represented as

$$A^* + h\nu \rightarrow A + 2h\nu$$

This kind of emission is responsible for laser action.

In this process, the probability of transition of atoms from E_2 to E_1 is depending on N_2 and U_ν

Therefore, the rate of stimulated emission is $\propto N_2U_\nu = B_{21}N_2U_\nu$

Where B_{21} is a proportionality constant known as the Einstein's coefficient for stimulated emission.

Expression for Energy density in terms of Einstein Coefficients:

Consider a system containing of large number of atoms is under thermal equilibrium. Let E_1 and E_2 be the lower and higher energy levels which contains N_1 and N_2 number of atoms respectively. Let U_ν be the energy density of the incident radiation. Hence the system absorbs and emits the energy through the processes of induced absorption spontaneous emission and stimulated emission. The energy of the photon absorbed and emitted by the atoms is $E = h\nu = E_2 - E_1$.

The system be in thermal equilibrium; the total energy of the system remains unchanged in spite of the interaction of the incident radiation.

At thermal equilibrium,

Rate of Absorption = Rate of Spontaneous emission + Rate of Stimulated emission

$\therefore$ we have,

$$B_{12}N_1U_\nu = A_{21}N_2 + B_{21}N_2U_\nu$$

$$B_{12}N_1U_\nu - B_{21}N_2U_\nu = A_{21}N_2$$

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

$$U_\nu = \frac{A_{21}N_2}{(B_{12}N_1 - B_{21}N_2)}$$

$\div$ both numerator & denominator of the above equation by B_{21} and N_2, we get,

$$U_\nu = \frac{A_{21}}{B_{21}}\left(\frac{1}{\frac{B_{12}N_1}{B_{21}N_2} - 1}\right) \qquad \rightarrow (4)$$

According to Boltzmann distribution law, we have

$$N_i = N_0 e^{\frac{-E_i}{KT}} \qquad \therefore \frac{N_1}{N_2} = \frac{e^{-\frac{E_1}{KT}}}{e^{-\frac{E_2}{KT}}} = e^{\frac{E_2 - E_1}{KT}} = e^{\frac{h\gamma}{KT}} \text{ OR } \frac{N_2}{N_1} == e^{-\frac{h\nu}{KT}} \qquad \rightarrow (5)$$

Now, by substituting this in equation (4), we have

$$U_\nu = \frac{A_{21}}{B_{21}}\left(\frac{1}{\frac{B_{12}}{B_{21}}e^{\frac{h\nu}{KT}} - 1}\right) \qquad \rightarrow (6)$$

According to Planck's law, the equation for energy density E_γ is given by

$$U_\nu = \frac{8\pi h\nu^3}{c^3}\left(\frac{1}{e^{\frac{h\nu}{KT}} - 1}\right) \qquad \rightarrow (7)$$

Now, comparing equation (6) and equation (7), we have

$$\frac{A_{21}}{B_{21}} = \frac{8\pi h\nu^3}{c^3} \;\&\; \frac{B_{12}}{B_{21}} = 1 \text{ or } B_{12} = B_{21}$$

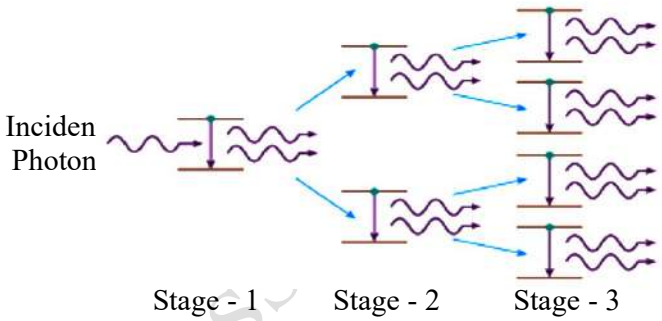
This implies that the probability of induced absorption is equal to the probability of stimulated emission. Because of the above identity, A_{21} and B_{21} can be represented by A and B in equation (6) and it can be rewrite as.

$$U_\nu = \frac{A}{B}\left(\frac{1}{e^{\frac{h\nu}{KT}} - 1}\right)$$

This is the expression for energy density in terms of Einstein's A and B coefficients.

Condition for Lasing action:

Consider a LASER system. Let an atom in the excited state is stimulated by a photon of suitable energy, so that atom makes stimulated emission. Then two coherent photons are obtained. These two coherent photons if stimulate two more atoms in the exited state to make emission then four coherent photons are obtained. These four coherent photons stimulate four more atoms in the excited state resulting in eight coherent photons and so on. As this process continues, number of coherent photons produced increases. These coherent photons constitute an intense beam of LASER. This phenomenon of building up of number of coherent photons, so as to get an intense LASER beam is called lasing action.



LASER action could be achieved through the conditions of population inversion by pumping and meta-stable state.

(i) Population Inversion:

When a system is under thermal equilibrium, the number of atoms in excited state is less than the number of atoms in the lower energy state. The production of LASER is achieved through stimulated emission rather than induced absorption and spontaneous emission. This is possible only if the number of atoms in the higher energy state is more than the number of atoms in the lower energy state and the process of achieving this is called population inversion.

Thus the essential conditions for population inversion are

a) Higher energy state should possess a longer life time.

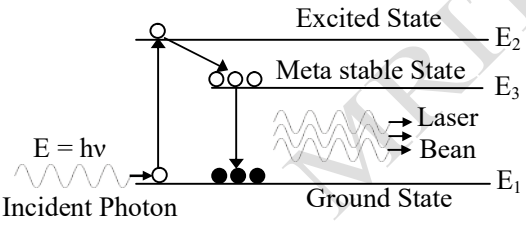
b) The number of atoms in the higher energy state must be greater than the number of atoms in the lower energy state.

(ii) Pumping Process:

Population inversion is achieved by supplying energy from a suitable source is called Pumping. In addition, to have more stimulated emissions, the life time of atoms in the excited state must be longer. There are number of techniques for pumping a collection of atom to an inverted state (excited state).

(iii) Meta stable state:

The life time of an energy level is of the order of 10^{-8} second. If an atom possesses unusual longer life time in an energy state such a state is referred to as a meta-stable state. Usually the life time of meta-stable state varies from 10^{-2}s to 10^{-3}s. Population inversion could be achieved with the help of three energy state with one of them a meta-stable state and is as shown in the figure. The population inversion is achieved between the state E_2 and E_1 as state E_3 is a meta-stable state.



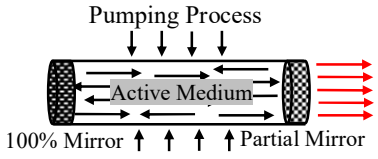
Note : The principles of Laser are 1. Stimulated Emission, 2. Population Inversion & 3. Meta-stable State

Requisites of a LASER system:

The three requisites of a LASER system are;

1. Active medium
2. Pumping Source
3. LASER cavity

1. Active medium: Population inversion occurs at certain stage in the active medium due to the absorption of energy. The active medium supports meta-stable states. After this stage the active medium is capable of emitting LASER light.



2. Pumping Source: Pumping source is an excitation source in order to achieve population inversion. That means more and more atoms are to be moved to higher energy state is called pumping. This is achieved by supplying suitable energy from an energy source. optical / light energy is used in optical pumping and electrical energy is used electrical pumping.

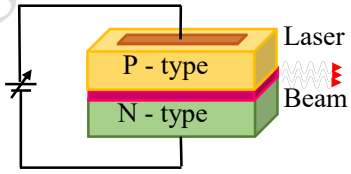
3. LASER cavity: The LASER Cavity is an active medium bound between two highly parallel mirrors. The reflection of photons from the mirrors results in multiple traverse of photons through the active medium inducing more and more stimulated emissions. Thus amplification of light is achieved. This also helps to tap certain permissible part of LASER energy from the active medium. The cavity resonates and the output will be maximum when the distance L between the mirrors is equal to an integral multiple of λ/2 . $i.e., \quad L = \frac{\lambda}{2}$

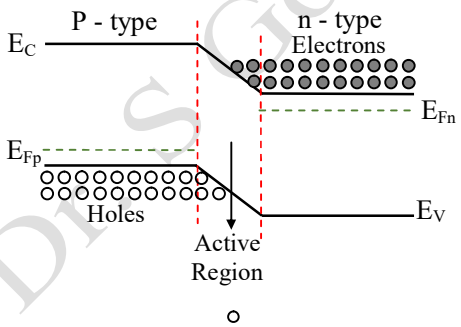
Where, λ is the wavelength of incident suitable radiation and L is the length of the LASER cavity.

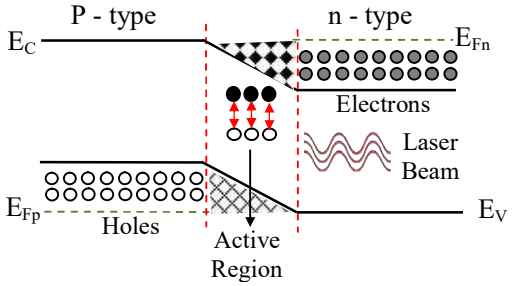
Semiconductor laser:

Principle: A semiconductor diode laser is a specially fabricated P-N junction device that emits coherent light when it is forward biased. It works on the principle of LED where p-type and n-type semiconductors are heavily doped. Hence electrons and holes are recombined in the depletion region producing coherent beam radiations.

Construction: The schematic diagram of semiconductor diode laser is shown in figure. It consists of a heavily doped n and p regions. The n-region is obtained by doping with pentavalent tellurium and the p-region is obtained by doping with trivalent zinc. The p-n region lies in a horizontal plane through the centre. The top and bottom faces of a diode are metalized to pass current through the diode. The front and rear faces are well polished parallel to each other and perpendicular to the plane of the junction. The other two opposite faces are roughened to prevent the lasing action in that direction. The first figure indicates energy level diagram for ordinary or LED diode and the second figure indicates the energy level diagram for Laser diode.







Working: The Diode is forward biased using an external source. Therefore, electrons and holes flow across the junction. The injected electrons and holes in the depletion region cause spontaneous emission of photons and the junction acts as LED. A population inversion is achieved in the depletion region of heavily doped P-N junction semiconductor diode in forward biased. Hence more electrons are occupied in donor levels and conduction band of n-type semiconductor, and the Fermi level lies

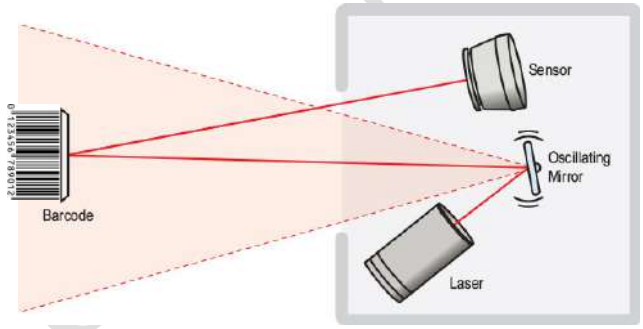
within the conduction band. Similarly, the acceptor levels are unoccupied and more holes are existing in the valence band and the Fermi level lies within the valence band.

When a diode is forward biased, the energy levels shifted and the new distribution is as shown in 2nd figure. As the current is increased, the intensity of light increased. When the current reaches a threshold value the carrier concentration in the depletion region will reach very high values. This region contains a large concentration of electrons in CB and holes in VB. i.e., upper levels in the depletion region are having high population while the lower levels are vacant. This is the population inversion. The narrow region where the state of population inversion is achieved is called active region. The Stimulated electron - hole recombination cause emission of coherent beam of radiation. At room temperature GaAs semiconductor diode emits laser light of wavelength 9000Å in IR region. A GaAsP emits laser of wavelength 6500Å in the visible region as red light.

The semiconductor diode lasers are simple, compact and highly efficient. They require very little power. Diode lasers give more divergent beam having an angular spread of the order of 5°-15°. They are less monochromatic and highly temperature sensitive. In semiconductor diode there is no meta-stable state.

Engineering applications of Lasers:

(1) Bar Code Scanner:



A bar code consists of a series of strips of white and black bars. Each strip has a width of about 0.3 mm and the total width of the bar code is about 3 cm. By simple scanning complete information about the product can be obtained. A typical scanning speed is about 200 m/s. In the bar code scanner, a low power (~ 0.5 mW) laser beam is reflected by a rotating polygon mirror to scan along a line. When the laser beam hits the bar code, the amount of reflected light from the bar code is captured and then the decoder converts the black and white bars with the binary signals. These reflected light signals (binary signals) are focused on a photo detector, which converts the light signal to an electrical signal. Further these electrical signals are then converted in to text data and are delivered to a computer by the scanner.

In order to be able to scan the product in any arbitrary direction for ease of scanning, the laser beam is made to scan in multiple directions by using multiple mirrors with the rotating polygon.

Advantages of laser barcode scanner:

i) The laser barcode scanner is used for non-contact scanning flexibly and efficiently.

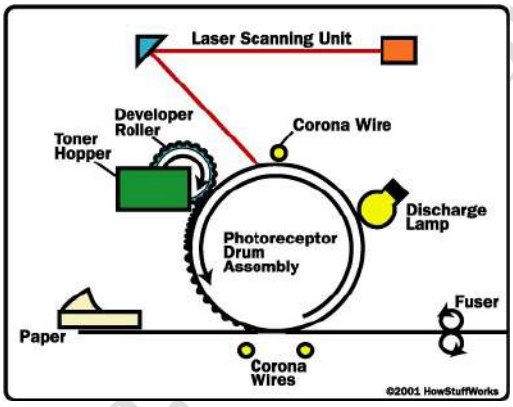
ii) The laser barcode scanner is the only choice when the scanning distance exceeds 30 cm.

iii) The laser barcode scanner can scan irregular barcode surface and read through glass or transparent adhesive paper, and it will not damage the barcode label.

iv) The Human errors are minimized and eliminated while using barcode scanners, therefore they are utilized and used in majority of the business globally.

v) The barcode readers are very accurate while collecting the data from sticker and labels therefore the chances of making errors are neutralized.

(2) Laser Printer:



Laser printer was invented at XEROX in 1969 by researcher Gary Starkweather. Laser printers are digital printing devices that are used to create high quality text and graphics on plain paper. A diode laser is used in the process of printing in Laser printer.

Construction: A laser printer consists of three main blocks, namely the scanning unit, toner cartridge unit and fuser assembly unit. The scanning unit of a laser printer typically consists of a laser diode, a scanning motor and a polygon mirror.

Working principles are:

1. A laser beam projects an image of the page to be printed onto an electrically charged rotating photo sensitive drum coated with selenium.
2. Photo conductivity allows charge to leak away from the areas which are exposed to light and the area gets positively charged.
3. Toner particles are then electrostatically picked up by the drum's charged areas, which have been exposed to light.
4. The drum then prints the image onto paper direct contact and heat, which fuses the ink to the paper.

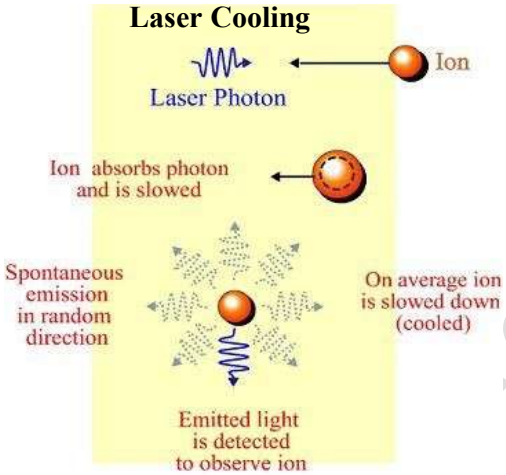
Advantages:

1. Laser printers are generally quiet and fast.
2. Laser printers can produce high quality output on ordinary papers.
3. The cost per page of toner cartridges is lower than other printers.

Disadvantages:

1. The initial cost of laser printers can be high.
2. Laser printers are more expensive than dot-matrix printers and ink-jet printers

(3) Laser Cooling:



Principle of Laser cooling is the use of dissipative light forces for reducing the random motion and thus the temperature of small particles, typically atoms or ions. Depending on the mechanism used, the temperature achieved can be in the millikelvin, microkelvin or even nanokelvin regime. Laser cooling is a technique used to slow down and manipulate the motion of atoms or molecules using laser light. By carefully tuning the frequency of the laser light to match the energy difference between atomic energy levels, it is possible to transfer momentum to the atoms and reduce their kinetic energy.

If an atom is traveling toward a laser beam and absorbs a photon from the laser, it will be slowed by the fact that the photon has momentum $P = mc = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}$. It would take a large number of such absorptions to cool the sodium atoms to near 0K. There are two types of laser cooling; they are Doppler Cooling and Sisyphus Cooling.

Problems:

(1) Show that the ratio of rate of spontaneous emission to induced absorption is given by $\left[1-e^{-\frac{h\nu}{kT}}\right]$

$$\frac{\text{Rate of spontaneous emission}}{\text{Rate of induced absorption}}=\frac{A_{21}N_2}{B_{12}N_1U(\nu)}$$

$$But,\frac{N_2}{N_1}=e^{-\frac{h\nu}{kT}} \quad \text{and} \quad U_\nu=\frac{A_{21}}{B_{21}}\left[\frac{1}{e^{\frac{h\nu}{kT}}-1}\right]$$

$$\therefore \frac{\text{Rate of spontaneous emission}}{\text{Rate of induced absorption}}=\frac{A_{21}}{B_{12}}e^{-\frac{h\nu}{kT}}\frac{1}{\frac{A_{21}}{B_{21}}\left[\frac{1}{e^{\frac{h\nu}{kT}}-1}\right]}$$

$$Since \quad B_{12}=B_{21}$$

$$\frac{\text{Rate of spontaneous emission}}{\text{Rate of Induced absorption}}=e^{-\frac{h\nu}{kT}}\left[e^{\frac{h\nu}{kT}}-1\right]=\left[1-e^{-\frac{h\nu}{kT}}\right]$$

(2) Find the number of modes and their frequency separation in a resonant cavity of length 1 m of a laser operating at wavelength 632.8nm.

Soln.; L = 1m, λ = 632.8 nm = 632.8x10^{-9}m, n =?

$$WKT; \quad \lambda=\frac{2L}{n} \quad \text{OR} \quad n=\frac{2L}{\lambda}$$

$$n=\frac{2\times1}{632.8\times10^{-9}}=3160556$$

$$Also, \quad \lambda=\frac{c}{\nu}=\frac{2L}{n} \quad \text{OR} \quad \nu=\frac{nc}{2L}$$

$$\therefore \nu_n=\frac{nc}{2L} \quad and \quad \nu_{n-1}=\frac{(n-1)c}{2L}$$

$$\therefore Frequency\ separation=\nu_n-\nu_{n-1}$$

$$\therefore Frequency\ separation=\frac{nc}{2L}-\frac{(n-1)c}{2L}=\frac{nc-nc+c}{2L}=\frac{c}{2L}$$

$$\therefore Frequency\ separation=\frac{c}{2L}=\frac{3\times10^8}{2\times1}=1.5\times10^8\,Hz$$

(3) A laser operating at 632.8nm emits 3.182x10^{16} photons per second. Calculate the output power of the laser. Also find the percentage power converted into coherent light energy, if the input power is 100 watt.

Soln.; λ = 632.8nm = 632.8x10^{-9}m, N = 3.182x10^{-16}photon/sec, p = ?, input power = 100 watt, %age power =?

$WKT, \quad N.\Delta E = E = Pt$

$\therefore N.\frac{hc}{\lambda} = Pt \quad \text{OR} \quad P = \frac{Nhc}{\lambda t}$

$P = \frac{3.182\times10^{16}\times6.626\times10^{-34}\times3\times10^{8}}{632.8\times10^{-9}x1} = 0.01\,watt$

$P = 0.01\,watt$

$Percentage\ Power = \frac{output\ power}{input\ power}\times100 = \frac{0.01}{100}\times100 = 0.01$

$Percentage\ Power = 0.01\%$

(4) Calculate the wavelength of laser emitted from an extrinsic semiconductor laser if the band gap is 0.02eV. To which region of spectrum does it belong?

Soln.; λ = ?, E_g = 0.02 eV = 0.02x1.602x10^{-19} J

$WKT, \quad E_g = h\nu = \frac{hc}{\lambda}$

$\therefore \lambda = \frac{hc}{E_g} = \frac{6.626\times10^{-34}\times3\times10^{8}}{0.02\times1.602\times10^{-19}} = 6.208\times10^{-5}\,m$

This wavelength region belongs to infrared region (0.01 to 7x10^{-5}m).

NOTE: Visible region 7x10^{-5}m to 4x10^{-5}m, ultraviolet 4x10^{-5}m to 10^{-7}m

(5) A pulse from laser with power 1mW lost for 10nS, if the number of photons emitted per pulse is 3.491x10^{7}. Calculate the wavelength of laser. (May22)

Soln.; P= 1 x10^{-3}W, t=10 x10^{-9}sec, N=3.491x10^{7}, λ=?

$N.\Delta E = Pt$

$N\frac{hc}{\lambda} = Pt \Rightarrow \lambda = \frac{Nhc}{Pt}$

$\therefore \lambda = \frac{3.491\times10^{7}\times6.625\times10^{-34}\times3\times10^{8}}{1\times10^{-3}\times10\times10^{-9}} = 6.938\times10^{-7}\,m$

(6) A pulsed laser emits photons of wavelength 780nm with 20mW average power per pulse. Calculate the number of photons contained in each pulse if the pulse duration is 10ns.

Soln.; λ=780 x10^{-9}m, P= 20 x10^{-3}W, N=?, t=10 x10^{-9}sec

$WKT, \quad N.\Delta E = Pt$

$N\frac{hc}{\lambda} = Pt \quad \text{OR} \quad N = \frac{\lambda Pt}{hc}$

$\therefore N = \frac{780\times10^{-9}\times20\times10^{-3}\times10\times10^{-9}}{6.625\times10^{-34}\times3\times10^{8}} = 7.849\times10^{8}\ photons$

(6a) A laser source has a power output of 10^{-3}W. Calculate the number of photons emitted per second given wavelength of laser 692.8nm. (Model QP)

Soln., $\lambda = 692.8 \times 10^{-9}$m, $P = 10^{-3}$W, N = ?, t = 1sec

$$WKT, \quad N = \frac{\lambda P t}{hc}$$

$$\therefore N = \frac{692.8\times10^{-9}\times10^{-3}\times1}{6.625\times10^{-34}\times3\times10^{8}} = 3.486\times10^{15} \text{ photons/second}$$

(7) In a laser system, operating at 323K, the wavelength of the light emitted is 1.3μm. Determine the ratio of population of the energy levels (Boltzmann's factor).

Soln., T= 323K, $\lambda = 1.3 \times 10^{-6}$m, $\frac{N_2}{N_1} = ?$

$$WKT, \quad \frac{N_2}{N_1} = e^{-\frac{h\nu}{KT}} = e^{-\frac{hc}{\lambda KT}}$$

$$\frac{N_2}{N_1} = e^{\frac{-6.626\times10^{-34}\times3\times10^{8}}{1.3\times10^{-6}\times1.38\times10^{-23}\times323}} = e^{-34.304} = 1.265x10^{-15}$$

(8) The ratio of population of two energy levels out of which one corresponds to meta stable state is 1.059×10^{-30}. Find the wavelength of light emitted at 330K.

Soln., $\frac{N_2}{N_1} = 1.059\times10^{-30}$, T = 330K, $\lambda = ?$

$$WKT, \quad \frac{N_2}{N_1} = e^{-\frac{h\nu}{KT}} = e^{-\frac{hc}{\lambda KT}}$$

$$OR \quad \ln\left(\frac{N_2}{N_1}\right) = \frac{-hc}{\lambda KT}$$

$$\therefore \lambda = \frac{-hc}{KT\ln\left(\frac{N_2}{N_1}\right)} = \frac{-6.626\times10^{-34}\times3\times10^{8}}{1.38\times10^{-23}\times330\times\ln\left(1.059\times10^{-30}\right)} = 6.323x10^{-7}\,m$$

$$OR \quad \lambda = 632.3nm$$

(8a) Calculate the ratio of population for a given pair of energy levels corresponding to emission of radiation 694.3nm at a temperature of 300K. (Model QP)

$$WKT, \quad \frac{N_2}{N_1} = e^{-\frac{h\nu}{KT}} = e^{-\frac{hc}{\lambda KT}}$$

$$\therefore \frac{N_2}{N_1} = e^{\frac{-6.626\times10^{-34}\times3\times10^{8}}{694.3\times10^{-9}\times1.38\times10^{-23}\times300}} = e^{-69.155} = 9.25\times10^{-31}$$

(9) Calculate the ratio of (i) Einstein's coefficients and (ii) stimulated to spontaneous emissions, for a system in thermal equilibrium at 300K in which radiations of wavelength 1.39µm are emitted.

Soln., $\frac{A_{21}}{B_{21}}=?\frac{A_{21}}{B_{21}}=?$ and $\frac{B_{12}}{A_{21}}=?$, Rate of stimulated to spontaneous=? λ=1.39 X10^{-6}m,

T=300K

$$(i)\,WKT,\quad \frac{A_{21}}{B_{21}}=\frac{8\pi h\nu^3}{c^3}=\frac{8\pi h}{\lambda^3}=\frac{8\pi\times 6.626\times10^{-34}}{\left(1.39\times10^{-6}\right)^3}=6.2\times10^{-15}$$

$$\frac{A_{21}}{B_{21}}=6.2\times10^{-15}$$

$$(ii)\,WKT,\ \frac{\text{Rate of stimulated emission}}{\text{Rate of spontaneous emission}}=\frac{B_{21}N_2U(\nu)}{A_{21}N_2}$$

$$But,\ U_\nu=\frac{A_{21}}{B_{21}}\left[\frac{1}{e^{\frac{h\nu}{kT}}-1}\right]$$

$$\therefore \frac{\text{Rate of stimulated emission}}{\text{Rate of spontaneous emission}}=\frac{B_{21}}{A_{21}}.\frac{A_{21}}{B_{21}}\frac{1}{\left(e^{\frac{hc}{\lambda kT}}-1\right)}=\frac{1}{\left(e^{\frac{hc}{\lambda kT}}-1\right)}$$

$$\therefore \frac{\text{Rate of stimulated emission}}{\text{Rate of spontaneous emission}}=\frac{1}{\left(e^{\frac{6.626\times10^{-34}\times3\times10^{8}}{1.39\times10^{-6}\times1.38\times10^{-23}\times300}}-1\right)}=\frac{1}{\left(e^{34.543}-1\right)}=\frac{1}{\left(1.004x10^{15}-1\right)}=9.96\times10^{-16}$$

Model Questions:

1. What is LASER? Enumerate the Characteristics of a LASER Beam.
2. Discuss the three possible ways through which radiation and matter interaction can take place.
3. Explain the terms, (i) Induced absorption, (ii) Spontaneous emission, (iii) Stimulated emission, (iv) Population inversion, (v) Meta-stable state & (vi) Resonant cavity.
4. Explain the rates of absorption and emission and hence derive an expression for energy density using Einstein's A and B coefficients.
5. Explain requisites of LASER system.
6. What is Semiconductor LASER? Describe with energy band diagram the construction & working of Semiconductor diode LASER along with applications.
7. Discuss the working of LASER barcode reader.
8. With the help of a sketch describe the principle, construction and working of the LASER Printer.
9. Explain LASER Cooling and its application.

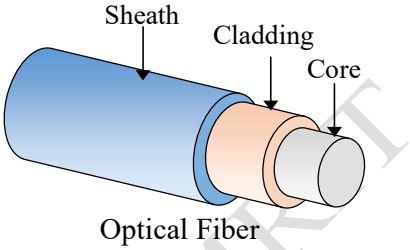
Numerical Problems:

1. Find the ratio of population of two energy levels in a LASER if the transition between them produces light of wavelength 6493 Å, assuming the ambient temperature at 27°C.
2. Find the ratio of population of two energy levels in a medium at thermal equilibrium, if the wavelength of light emitted at 291 K is 6928 Å.
3. The ratio of population of two energy levels out of which one corresponds to metastable state is 1.059×10^{-30}. Find the wavelength of light emitted at 330 K.
4. Find the ratio of population of two energy levels in a medium at thermal equilibrium, if the wavelength of light emitted at 300 K is 10μm. Also find the effective temperature when energy levels are equally populated.
5. The average power output of a LASER beam of wave length 6500 Å is 10 mW. Find the number of photons emitted per second by the LASER source.
6. The average power of a LASER beam of wavelength 6328 Å is 5 mW. Find the number of photons emitted per second by the LASER source.
7. A pulsed LASER has an average power output 1.5 mW per pulse and pulse duration is 20 ns. The number of photons emitted per pulse is estimated to be 1.047×10^{8}. Find the wavelength of the emitted LASER.
8. A pulsed LASER with power 1 mW lasts for 10 ns. If the number of photons emitted per pulse is 5×10^{7}. Calculate the wavelength of LASER.
9. A Ruby LASER emits a pulse of 20 ns duration with average power per pulse being 100 kW. If the number of photons in each pulse is 6.981×10^{15}, calculate the wavelength of photons.
10. In a LASER system when the energy difference between two energy levels is 2×10^{-19} J, the average power output of LASER beam is found to be 4 mW. Calculate number of photons emitted per second.

OPTICAL FIBERS

Introduction

Optical fiber is a device used to transmit light through bundle of thin fibers of transparent dielectric material from one end to another end for a very long distance. It works on the principle of Total Internal Reflection (TIR).



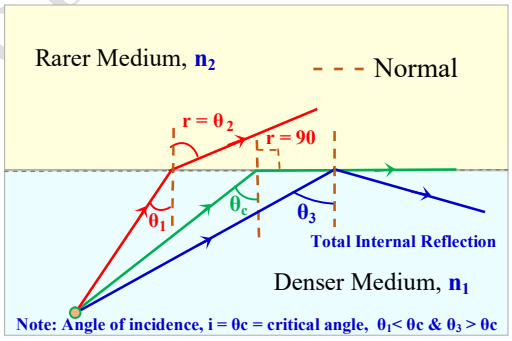
Construction: The sectional view of a typical optical fiber is as shown in the figure. It has three regions named as Core, Cladding and Sheath.

1. The innermost light guiding region is called Core.
2. The layer covering core is called Cladding or Clad, which helps in total internal reflection of light.
3. The outermost protective layer is called Sheath (Coating), which protects the fiber from mechanical stress and chemical reactions.

The optical fiber is designed to support total internal reflection and hence the refractive index (RI) of core n_1 is made greater than the RI of cladding n_2. A typical fiber will be of the order of few microns.

Total Internal Reflection

Consider a ray of light moving from a denser medium of refractive index n_1 to rarer medium of refractive index n_2. As a result, the incident ray of light bends away from the normal. Hence the angle of refraction θ_2 is greater than the angle of incidence θ_1. As the angle of incidence increases the angle of refraction also increases. For a particular angle of incidence, $i = \theta_c$ the refracted ray grazes the interface separating the two media. The corresponding angle of incidence θ_c is called **critical angle**. If the angle of incidence is greater than the critical angle θ_c, then the light ray is turned back into the same medium and is called **T**otal **I**nternal **R**eflection.



The above figure shows Total Internal Reflection;

According to Snell's law

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

when $\theta_1 = \theta_c$, then $\theta_2 = 90^\circ$ and $\sin 90^\circ = 1$

$$\therefore \quad n_1 \sin\theta_c = n_2$$

$$\therefore \quad \sin\theta_c = \frac{n_2}{n_1} \qquad \text{or} \qquad \theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)$$

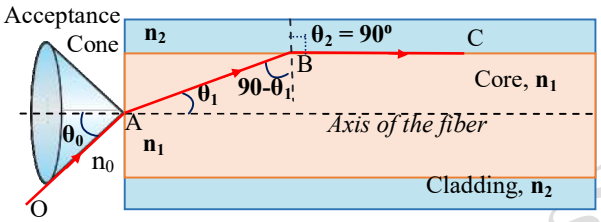
Angle of acceptance and Numerical aperture (NA):

Acceptance angle (θ) is the maximum angle of incidence with which the ray is sent into the fiber core which allows the incident light to be guided by the core. It is also called as waveguide acceptance angle or acceptance cone half angle.

In optics, the numerical aperture (NA) of an optical fiber is a dimensionless number that characterizes the range of angles over which the fiber can accept light. Numerical aperture represents the light gathering capability of optical fiber and it is defined as sine of the acceptance angle.

$$\therefore \text{NA} = \sin\theta_0$$

Condition for propagation (Derivation for Angle of Acceptance)



Consider an optical fiber with core made of refractive index n_1 and cladding made of refractive index n_2. Let n_0 be the refractive index of the surrounding medium.

Let a ray of light OA entering into the core at an angle of incidence θ_0 w.r.t fiber axis. Then it is refracted along AB at an angle θ_1 and meet the core-cladding interface at critical angle of incidence ($\theta_c = 90 - \theta_1$). Then the refracted ray grazes along BC.

By applying Snell's law at A, we get

$$n_0 \sin\theta_0 = n_1 \sin\theta_1$$

$$\sin\theta_0 = \frac{n_1}{n_0}\sin\theta_1 \qquad (1)$$

Again by applying Snell's law at B, we get

$$n_1 \sin(90-\theta_1) = n_2 \sin 90° \qquad \text{since, } \sin(90-\theta_1) = \cos\theta_1 \text{ and } \sin 90° = 1, \text{ we get}$$

$$n_1 \cos\theta_1 = n_2 \quad \text{or} \quad \cos\theta_1 = \frac{n_2}{n_1} \qquad (2)$$

$$\sin^2\theta_1 + \cos^2\theta_1 = 1 \quad \text{or} \quad \sin\theta_1 = \sqrt{1-\cos^2\theta_1}$$

$$\therefore \sin\theta_1 = \sqrt{1-\frac{n_2^2}{n_1^2}} \qquad (3)$$

By substituting this in equation (1), we get

$$\sin\theta_0 = \frac{n_1}{n_0}\sqrt{1-\frac{n_2^2}{n_1^2}} = \frac{n_1}{n_0}\frac{1}{n_1}\sqrt{n_1^2-n_2^2} = \frac{1}{n_0}\sqrt{n_1^2-n_2^2} \qquad (4)$$

Since, Numerical Aperture, NA = $\sin\theta_0$,

$$\therefore NA = \frac{1}{n_0}\sqrt{n_1^2-n_2^2}$$

If the surrounding medium is air, then $n_0 = 1$

$$\therefore NA = \sin\theta_0 = \sqrt{n_1^2 - n_2^2} \qquad (5)$$

This is the expression for Numerical Aperture of an optical fiber.
Light is transmitted through the fiber only when

$$\therefore \theta_i \le \theta_0 \quad \text{or} \quad \sin\theta_i \le \sin\theta_0 \le \sqrt{n_1^2 - n_2^2}$$

$$\therefore \sin\theta_i \le NA \le \sqrt{n_1^2 - n_2^2} \qquad (6)$$

This is the condition for propagation for light through the optical fiber with multiple total internal reflections.

From equation (4), we have

$$\sin\theta_0 = \frac{1}{n_0}\sqrt{n_1^2 - n_2^2}$$

$$\text{Or } \theta_0 = \sin^{-1}\left(\frac{1}{n_0}\sqrt{n_1^2 - n_2^2}\right)$$

If the surrounding medium is air, then $n_0 = 1$

$$\therefore \theta_0 = \sin^{-1}\left(\sqrt{n_1^2 - n_2^2}\right) = \sin^{-1}(NA)$$

This is the expression for Angle of Acceptance of an optical fiber.

Fractional index change (Δ):

It is the ratio of the difference between the refractive indices of the core and cladding to the refractive index of core of an optical fiber.

$$\text{i.e., } \Delta = \frac{n_1 - n_2}{n_1}$$

Δ is always positive and less than 1. Because, $n_1 > n_2$.

Relation between NA and Δ:

$$\text{WKT, } NA = \sqrt{n_1^2 - n_2^2} \quad \text{and} \quad \Delta = \frac{n_1 - n_2}{n_1}$$

$$NA = \sqrt{(n_1 - n_2)(n_1 + n_2)}$$

$$NA = \sqrt{\Delta n_1 (n_1 + n_2)}$$

$$\text{Since } n_1 \approx n_2, \; NA = \sqrt{\Delta n_1 (2n_1)}$$

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

Modes of Propagation: In an optical fiber the wave propagation mode is referred to as fiber modes. The light ray paths along which the waves are in phase inside the fiber are known as modes.

In simple words, the allowed paths for the light ray inside the fiber are known as modes of propagation.

$$\textbf{Number of modes} \propto \frac{\mathbf{d}}{\boldsymbol{\lambda}}$$

Where, d is the diameter of the core and λ is the wavelength of the light ray travelling through the fiber.

The different types of fiber modes are guided mode, leaky mode and radiation mode. In an optical communication system, signals are transmitted using guided modes only.

V- Number: The number of modes supported for propagation in the fiber depends on the core size, operating wavelength and refractive indices of core and cladding materials. It can be conveniently defined using a normalized frequency parameter, called V- number and it is given by

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

Where λ is the (operating) wavelength of light propagating in the fiber.

d is the core diameter

n_1 is the refractive index of the core.

n_2 is the refractive index of the cladding.

For V>>1, the number of modes supported by the fiber is given by

$$N = \frac{V^2}{2}$$

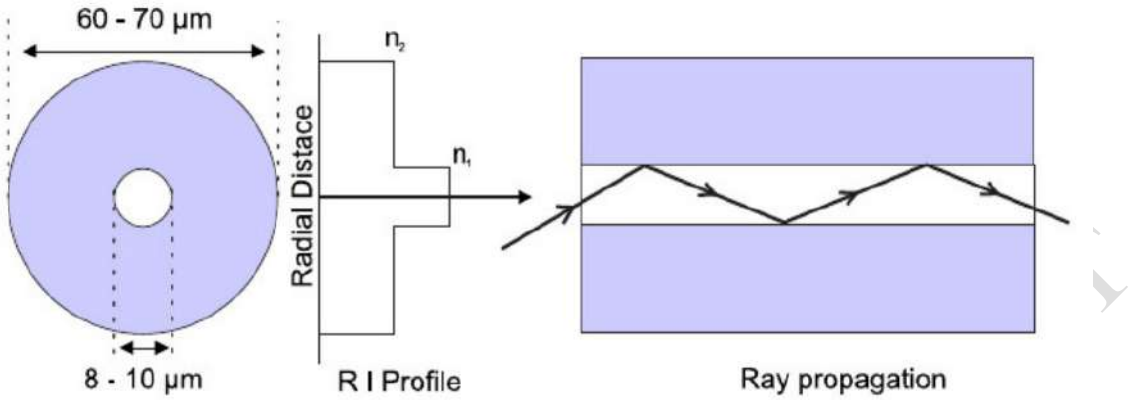
Refractive index profile: Refractive index profile is the distribution of refractive indices of materials within an optical fiber. Some optical fiber has a step-index profile, in which the core has one uniformly-distributed index and the cladding has a lower uniformly-distributed index. Other optical fiber has a graded-index profile, in which the refractive index varies gradually as a function of radial distance from the fiber center.

Types of optical fiber:

Based on the refractive index profile, core size and mode of propagation, the fibers are classified into three types, namely

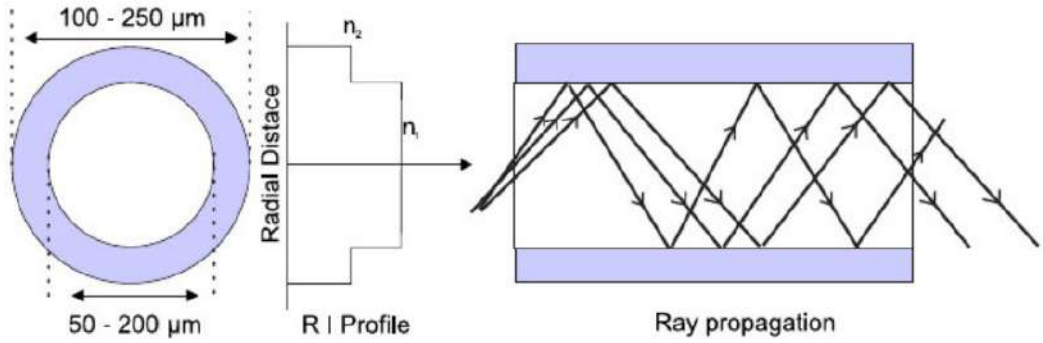
1. Step-index single mode fiber
2. Step-index multi-mode fiber
3. Graded-index multi-mode fiber

1. Step-index single mode fiber:



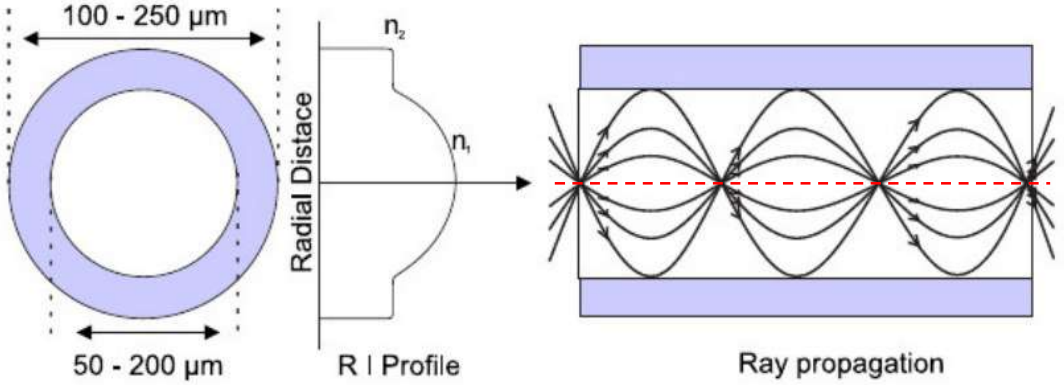
A single mode step index fiber consists of a very fine thin core (made of glass material) of uniform refractive index n_1 is surrounded by a cladding of refractive index n_2 lower than that of the core, Since there is abrupt change in RI of core and cladding at the interface it is called step index fiber. The diameter of the core is about 8 to 10 µm and that of cladding is about 60 to 70 µm. Since the core size is small, the numerical aperture is also small and hence it supports for single mode as shown in the figure. This need laser as the source of light.

2. Step-index multi-mode fiber:



Step-index multimode fiber is similar to that of a single mode fiber, but it has a larger core diameter and constant refractive index, by the virtue of which it will be able to support for large number of modes of propagation as shown in figure. The diameter of the core is about 50 to 200 µm and that of the cladding is about 100 to 250 µm. The step-index multi mode fiber can accept either a laser or LED as source of light. It is the least expensive of all. They are used in data links.

3. Graded-index multi-mode fiber:



A graded multimode fiber has concentric layers of RI is called GRIN fiber. That means the RI of the core varies with distance from the axis of the fiber. The refractive index of the core along the axis of the fiber is maximum and it decreases uniformly on either side of the axis towards the core-cladding interface. Hence the refractive index profile follows a parabolic shape and light transmissions in the fiber are shown in figure. The diameter of the core and that of cladding are almost same as that of multi-mode index fibers. Either a laser or LED's are used as light sources for operating the system. It is the most expensive of all and used in telecommunication links.

Attenuation (Transmission loss or fiber loss):

Attenuation is the reduction in power or intensity of light as it travels in the fiber. The reduction may be due to light absorption, scattering and radiation losses (extensive fiber bends). The net attenuation is given by a factor called the attenuation coefficient (α) in dB/km and it is defined as optical power output to the optical power input for a fiber of length L and a wavelength of the propagating light.

$$\alpha = \frac{-10}{L}\log\left(\frac{P_{out}}{P_{in}}\right) \quad dB/Km$$

Attenuation in an optical fiber is comparatively less than that in coaxial cables.

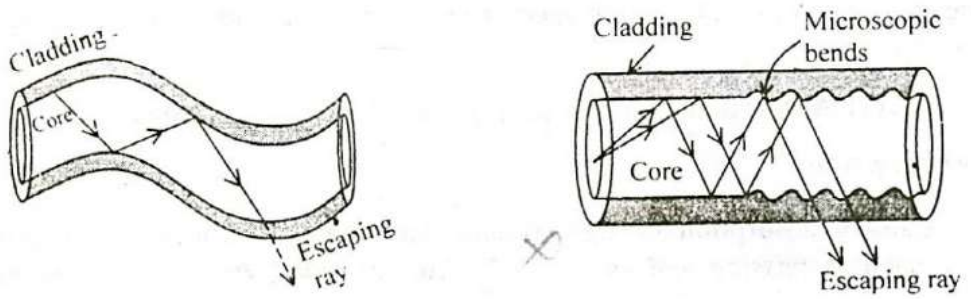
Causes of attenuation: The three mechanism through which attenuation takes place are

(1) Absorption losses: The absorption losses occur due to the presence of impurities (like Cr, Cu, ions trapped within the glass at the time of manufacture) or due to the basic material (glass) itself absorb energy at certain wavelength. Typical absorption losses are of the order of 0.1dB/km in the 0.8 to 1.6 µm wavelength ranges and 0.03dB/km in the 1.3 to 1.6 µm wavelength range.

(2) Scattering losses: Scattering losses occur due to imperfections and impurities in the fiber material. Refractive index changes while the signal travels in the fiber. This sharp variation in refractive index is induced by the localized structural inhomogeneity. This type of scattering is same as Rayleigh

scattering. Rayleigh scattering occurs whenever a light wave travels through a medium having scattering objects whose dimensions are smaller than a wavelength. Thus it becomes a loss.

(3) Geometrical loss (Radiation loss):



Macroscopic Bending

Geometrical losses occur due to (a) macroscopic and (b) microscopic bends.

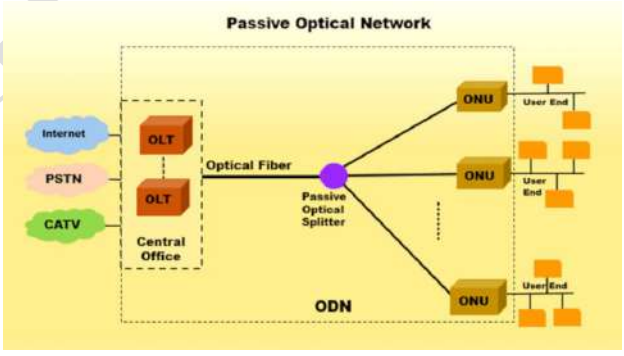
(a) Due to sharp bends, some of the light energy escapes through the cladding and leads to loss in the intensity of light ray.

(b) The microscopic bends cause irregular reflections and some of them then leak through the fiber.

Applications:

(1) Fiber Optic Networking:

A Local Area Network (LAN) is a type of computer net- work that interconnects multiple computers and computer-driven devices in a particular physical location. Traditionally copper coaxial cables are used for LAN.

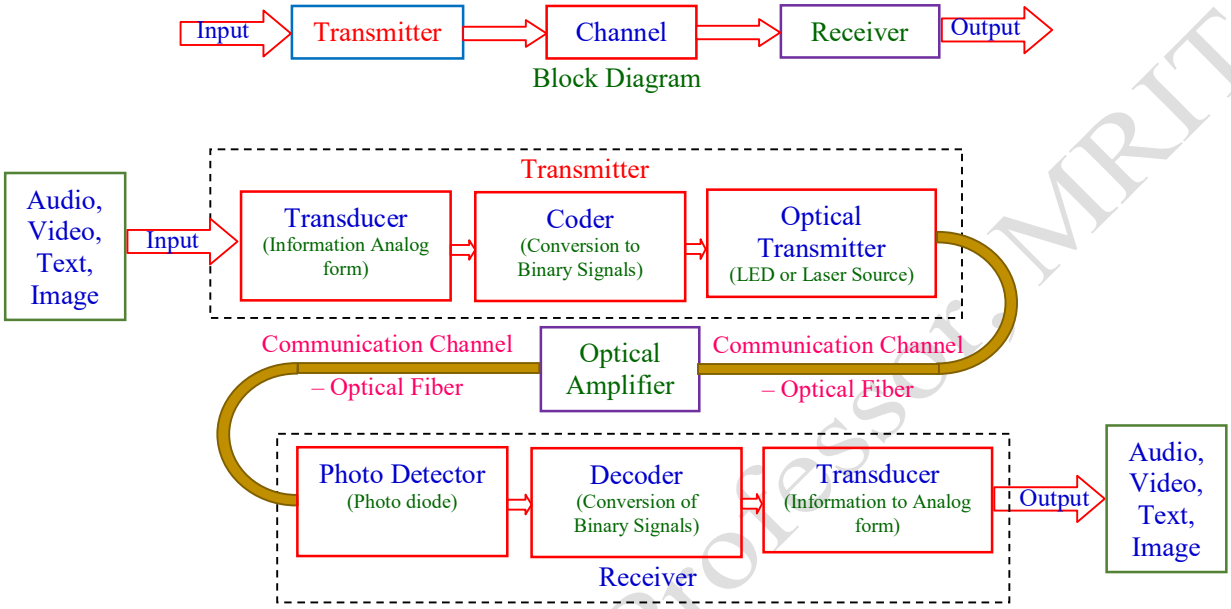


Passive optical LANs are built entirely using Optical fiber cables. The passive optical LAN is complicated as it works on the concept of optical network terminals (ONT) and passive optical splitters. Network switches act as passive splitters and the commercial media converters act as optical network terminals in a real-time application of passive opticalLAN.

Advantages:

1. High speeds and bandwidth
2. Longer distances are possible
3. Less chance of errors

(2) Fiber Optic Communication:



A basic of point to point communication system using optical fiber is shown in figure. The voice of a telephone user gives rise to electrical signals. These electrical signals are converted to binary data using coder. These electrical pulses are converted into pulses of optical power by an optical source (such as an LED or laser) in the binary form.

Now the light pulses is coupled into the optical fiber at an incidence angle less than that of the acceptance angle. The light pulses inside the fiber undergo total internal reflection and reach the other end of the fiber and fed into a light detector. Light detector converts the light signals into pulses of electric signals (current). These pulses are further decoded into analog electrical signal and converted into the usable form like audio or video.

Problems:

(1) Calculate the numerical aperture and angle of acceptance of a given optical fiber if the refractive index of the core and cladding are 1.563 and 1.498 respectively.

Soln.: NA = ?, θ_a = ? n_1 = 1.563 & n_2 = 1.498

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

$$NA = \sqrt{(1.563)^2 - (1.498)^2}$$

$$NA = 0.446$$

$$Sin\theta_a = NA$$

$$\theta_a = Sin^{-1}(0.446)$$

$$\theta_a = 26.5^0$$

(2) An optic glass fiber of refractive index 1.450 is to be clad with another glass to ensure internal reflection that will contain light traveling within 5 degree of the fiber axis. What maximum index of refraction is allowed for the cladding?

Soln.: n_1 = 1.450, $\theta_r = 5^0$ (or i = 90-5 = 85^0), n_2 = ?

$$\cos\theta_r = \frac{n_2}{n_1} \quad \text{or} \quad n_2 = n_1 \cos\theta_r$$

$$\therefore n_2 = 1.45\cos 5 = 1.444$$

(3) An optic fiber has a NA of 0.2 and a cladding refractive index of 1.59. Determine the core refractive index and also the acceptance angle for the fiber in water which has a refractive index of 1.33

Soln.: NA=0.2, n_2=1.59, n_1=? , θ_a=? , n_0=1.33

$$Sin\theta_a = NA$$

$$\therefore \theta_a = Sin^{-1}(0.2) = 11.54^0 = 11^\circ 32'$$

$$NA = \sqrt{\frac{n_1^2 - n_2^2}{n_0^2}}$$

$$n_1 = \sqrt{(NA)n_0^2 + n_2^2} = 1.612$$

(4) An optical has core refractive index 1.5 and clad refractive index 3% less than that of core. Calculate NA, angle of acceptance and internal critical angle.

Soln.: n_1 = 1.5, n_2 = 1.5 - (3% of 1.5), NA = ?, θ_a = ? & θ_c = ?

n_2 =1.5 - (0.03 x 1.5) =1.455

$$NA = \sqrt{n_1^2 - n_2^2} = \sqrt{(1.5)^2 - (1.455)^2} = 0.3647$$

$$\theta_a = Sin^{-1}(NA) = Sin^{-1}(0.3647) = 21.39^\circ$$

$$\theta_a = Sin^{-1}\left(\frac{n_2}{n_1}\right) = Sin^{-1}\left(\frac{1.455}{1.5}\right) = 75.93^\circ$$

(5) (Isem 2023) The angle of accepatance of an optical fibre is 30^0 when kept in air. Find the angle of accptance when it is in medium of refractive index 1.33

Soln.: Soln.: $\theta_a = 30°$, $n_0 = 1.33$, $\theta'_a = ?$

$$Sin\theta_a = \sqrt{n_1^2 - n_2^2} \quad \text{and} \quad Sin\theta_a^1 = \sqrt{\frac{n_1^2 - n_2^2}{n_0^2}}$$

$$\therefore \frac{Sin\theta_a^1}{Sin\theta_a} = \sqrt{\frac{n_1^2 - n_2^2}{n_0^2}} \frac{1}{\sqrt{n_1^2 - n_2^2}}$$

$$\frac{Sin\theta_a^1}{Sin\theta_a} = \frac{1}{n_0} \quad \text{or} \quad Sin\theta_a^1 = \frac{Sin\theta_a}{n_0}$$

$$Sin\theta_a^1 = \frac{Sin30}{1.33} \quad \text{or} \quad \theta_a^1 = Sin^{-1}[0.3759]$$

$$\theta_a^1 = 22.08^\circ$$

(6) A fiber sample 500 m long has an input power of 8.6 micro watt and an output power of 7.5 µW. What is the loss specification for the cable sample?

Soln.: L = 500 m = 0.5 km, P_{in} = 8.6 µW, P_{out} = 7.5 µW, $\alpha = ?$

$$\alpha = \frac{-10}{L}\log\left(\frac{P_{out}}{P_{in}}\right) dB/km$$

$$\alpha = \frac{-10}{0.5}\log\left(\frac{7.5}{8.6}\right) = 1.19\ dB/km,$$

(7) The attenuation of light in an optical fiber is estimated to be 2.0dB/km. What fraction of the initial intensity remains after 1 km and after 8 km?

Soln.: $\propto = 2.0$ dB/km, $\left(\frac{P_{out}}{P_{in}}\right)_{1Km} = ? \left(\frac{P_{out}}{P_{in}}\right)_{8Km} = ?$

$$\alpha = \frac{-10}{L}\log\left(\frac{P_{out}}{P_{in}}\right) dB/km \quad \text{or} \quad \left(\frac{P_{out}}{P_{in}}\right) = 10^{\left(\frac{-\alpha L}{10}\right)}$$

Intensity remaining after 1 km is

$$\left(\frac{P_{out}}{P_{in}}\right)_{1km} = 10^{\left(\frac{-2\times1}{10}\right)} = 0.631$$

Intensity remaining after 8 km is

$$\left(\frac{P_{out}}{P_{in}}\right)_{8km} = 10^{\left(\frac{-2\times 8}{10}\right)} = 0.025$$

(8) An optical signal propagating in a fiber retains 85% of input power after traveling a distance of 500 m in the fiber. Calculate the attenuation coefficient.

Soln.:

$$\alpha = \frac{-10}{L}\log\left(\frac{P_{out}}{P_{in}}\right) dB/km$$

$$\alpha = \frac{-10}{L}\log\left(\frac{85}{100}\right) = 1.412\ dB/km$$

(8) In a step index fiber, the relative refractive index difference is 2% and refractive index of cladding is 1.4. Calculate the refractive index of core and also the critical propagation angle.

Soln.: $\Delta=2\%=0.02$, $n_2=1.4$, $n_1=?$, $\theta_r=?$

$$\Delta = \frac{n_1 - n_2}{n_1} = 1 - \frac{n_2}{n_1}$$

$$or \quad \frac{n_2}{n_1} = 1-\Delta \quad or \quad n_1 = \frac{n_2}{1-\Delta}$$

$$\therefore \quad n_1 = \frac{1.4}{1-0.02} = 1.4285$$

$$\cos\theta_r = \frac{n_2}{n_1} = \frac{1.4}{1.4285} = 0.98$$

$$\theta_r = \cos^{-1}(0.98) = 11.48 = 11°29'$$

(9) Consider a slab waveguide made of AlGaAs having RI for core and clad 3.6 and 3.55 respectively. Find, how many modes can propagate in this waveguide if d = 5λ

Soln.: $n_1=3.6$, $n2=3.55$, $N=?$, $d=5\lambda$

$$N = \frac{V^2}{2} = \frac{1}{2}\left[\frac{\pi d}{\lambda}\sqrt{n_1^2 - n_2^2}\right]^2$$

$$N = \frac{1}{2}\left[\frac{\pi\times 5\lambda}{\lambda}\sqrt{(3.6)^2 - (3.55)^2}\right]^2$$

$$N = 44.10 \approx 44$$

(10) A step- index optical fiber has a core index of 1.46 and the cladding index of 1.409. If the core diameter is 80micro m and the wavelength of the light source is 1.2micro m, determine the number of modes present in the fiber.

Soln.: $n_1 = 1.46$, $n_2 = 1.409$, $d = 80\text{x}10^{-6}$ m, $\lambda = 1.2\text{x}10^{-6}$ m, $N = ?$

$$N = \frac{1}{2}\left[\frac{\pi d}{\lambda}\sqrt{n_1^2 - n_2^2}\right]^2$$

$$N = \frac{1}{2}\left[\frac{\pi \times 80\times10^{-6}}{1.2\times10^{-6}}\sqrt{(1.46)^2 - (1.409)^2}\right]^2$$

$$N = 3209.13 \approx 3209$$

(11) A single mode step index optical fiber used in communication has a core with refractive index 1.45, refractive index change of 5x10-3 and a core diameter of 6µm. If the operating wavelength of the communication system is 1.2 micro m, determine the V- parameter of the cable.

Soln.: n_1=1.45, $\Delta = 5\text{x}10^{-3}$, $d = 6\mu m = 6\text{x}10^{-6}$m, $\lambda = 1.2\text{x}10^{-6}$ m, $V = ?$

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

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

$$V = \frac{\pi \times 6\times10^{-6}}{1.2\times10^{-6}} \times 1.45 \times \sqrt{2\times5\times10^{-3}}$$

$$V = 2.28$$

(12) (May22) Calculate NA, Relative RI, V- number and the number of modes in an optical fiber of core diameter 50µm and core and cladding RI are 1.41 and 1.40 respectively. Given wavelength of source 820 nm.

Soln.: NA = ?, $\Delta = ?$, V = ?, N = ?, $d = 50\ \mu m = 50\text{x}10^{-6}$ m, $n_1 = 1.41$, $n_2 = 1.40$, $\lambda = 820\text{x}10^{-9}$ m

$$NA = \sqrt{n_1^2 - n_2^2} = \sqrt{(1.41)^2 - (1.40)^2} = 0.168$$

$$\Delta = \frac{n_1 - n_2}{n_1} = \frac{1.41 - 1.40}{1.41} = 7.09\times10^{-3}$$

$$V = \frac{\pi d}{\lambda}\sqrt{n_1^2 - n_2^2} = \frac{\pi \times 50\times10^{-6} \times \sqrt{(1.41)^2 - (1.40)^2}}{820\times10^{-9}} = 32.11 \qquad N = \frac{V^2}{2} = \frac{(32.11)^2}{2} = 515$$

(13) The refractive indices of the core and cladding of a step- index optical fiber are 1.45 and 1.40 respectively and its core diameter is 45µm. Calculate its relative refractive index difference, NA, V- number at wavelength 1000nm and the number of modes.

Soln.: $n_1 = 1.45$, $n_2 = 1.40$, $d = 45\ \mu m = 45\text{x}10^{-6}$ m, λ=1000 x10^{-9} m, $\Delta = ?$, NA = ?, V = ?, N =?

$$\Delta = \frac{n_1 - n_2}{n_1} = \frac{1.45 - 1.40}{1.45} = 0.0345$$

$$NA = \sqrt{n_1^2 - n_2^2} = \sqrt{(1.45)^2 - (1.40)^2} = 0.38$$

$$V = \frac{\pi d}{\lambda}\sqrt{n_1^2 - n_2^2} = \frac{\pi \times 45 \times 10^{-6} \times 0.38}{1000 \times 10^{-9}} = 53.72$$

$$N = \frac{V^2}{2} = \frac{(53.72)^2}{2} = 1443$$

Model Questions:

1. Define the terms: (i) angle of acceptance, (ii) numerical aperture, (iii) modes of propagation & (iv) refractive index profile.
2. Obtain an expression for numerical aperture and arrive at the condition for propagation.
3. Explain modes of propagation and RI profile.
4. What is attenuation? Explain the factors contributing to the fiber loss.
5. Discuss the types of optical fibers based on modes of propagation and RI profile.
6. Explain attenuation along with the expression for attenuation co-efficient and also discuss the types of fiber losses.
7. Explain the Fiber Optic Networking and mention its advantages.
8. Discuss point to point optical fiber communication system and mention its advantages over the conventional communication system.
9. Discuss the advantages and disadvantages of an optical communication.

Numerical Problems:

1. Calculate the numerical aperture and angle of acceptance for an optical fiber having refractive indices 1.563 and 1.498 for core and cladding respectively.
2. The refractive indices of the core and cladding of a step index optical fiber are 1.45 and 1.4 respectively and its core diameter is 45 μm. Calculate its fractional refractive index change and numerical aperture.
3. Calculate numerical aperture, acceptance angle and critical angle of a fiber having a core RI 1.50 and cladding RI 1.45.
4. An optical fiber has a numerical aperture of 0.32. The refractive index of cladding is 1.48. Calculate the refractive index of the core, the acceptance angle of the fiber and the fractional index change.
5. An optical signal propagating in a fiber retains 85% of input power after travelling a distance of 500 m in the fiber. Calculate the attenuation coefficient.
6. An optical fiber has core RI 1.5 and RI of cladding is 3% less than the core index. Calculate the numerical aperture, angle of acceptance critical angle.

7. The numerical aperture of an optical fiber is 0.2 when surround by air. Determine the RI of its core, given the RI of the cladding is 1.59. Also find the acceptance angle when the fiber is in water of RI 1.33.
8. The angle of acceptance of an optical fiber is 300 when kept in air. Find the angle of acceptance when it is in medium of refractive index 1.33.
9. An optical fiber of 600 m long has input power of 120 mW which emerges out with power of 90 mW. Find attenuation in fiber.
10. The attenuation of light in an optical fiber is 3.6 dB/km. What fraction of its initial intensity is remains after i) 1 km and ii) 3 km ?
11. The attenuation of light in an optical fiber is 2.2 dB/km. What fraction of its initial intensity is remains after i) 2 km and ii) 6 km?

MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2022-23

Module – 2 : Quantum Mechanics

Notes

Syllabus:

Quantum Mechanics: **8 hrs**

deBroglie Hypothesis and Matter Waves, deBroglie wavelength and derivation of expression by analogy, Phase Velocity and Group Velocity, Heisenberg's Uncertainty Principle and its application (Non - existence of electron inside the nucleus - Non Relativistic), Principle of Complementarity.

Wave Function, Time independent Schrödinger wave equation (Derivation), Physical Significance of a wave function and Born Interpretation, Expectation value, Eigen functions and Eigen Values, Particle inside one dimensional infinite potential well, Quantization of Energy States, Waveforms and Probabilities. Numerical Problems.

Pre requisite: Wave–Particle dualism

Self-learning: de Broglie Hypothesis

Quantum Mechanics

The branch of mechanics that deals with the mathematical description of the motion and interaction of subatomic particles, incorporating the concepts of quantisation of energy, wave – particle duality, the uncertainty principle and the corresponding principles.

Introduction:

At the end of 19^{th} century and in the beginning of 20^{th} century, many new phenomena such as photoelectric effect, Compton Effect, pair production, Zeeman Effect, radiation effects, nuclear radiations etc., were discovered. Since classical mechanics fails to explain the above phenomena, a new physics known as modern physics was developed on the basis of quantum theory of radiation. In order to explain the distribution of energy in the blackbody radiation Planck introduced the concept of quantum theory of radiation in 1900.

Radiation:

Radiation is defined as the energy that travels through space or matter in the form of a particle or wave.

In physics, radiation is the emission or transmission of energy in the form of waves or particles through space or a material medium.

It is a process of transmission of energy from one place to another without the aid of any intervening medium.

Light and heat are types of radiation. Heat radiation is also called thermal radiation.

OR

"Radiation is the emission of electromagnetic waves by matter when supplied with appropriate amount of energy".

A photon is the smallest discrete packet or quantum of electromagnetic radiation. It is the basic unit of all light.

The basic properties of photons are:

- A packet or bundle of energy is called a photon
- The energy of a photon is $E = h\nu = hc/\lambda$
- The momentum of the photon is $p = E/c = h/\lambda$
- Photon can carry energy and momentum which are dependent on the frequency.
- The rest mass of the photon is zero and hence they can exist only moving states.
- They are charge less particles and are not affected by either electric or magnetic field,
- They are stable and having integral spin (spin-1 particles) which make them as bosons.
- They can have interactions with other particles such as electrons, protons, neutrons, etc.,
- They can travel with the speed of light in free space or vacuum.

Planck's Quantum Theory of Radiation:

In the year 1900, Max Planck propounded the quantum theory of radiation. According to this theory, the emission and absorption of radiation is not continuous but it is in the form of packets of definite and discrete set of energy. Each packet is called the quanta or photon which has a definite energy and definite momentum whose value is proportional to the frequency of radiation. If 'ν' is the frequency of radiation, the energy of each photon is given by,

$$E \propto \nu \qquad \text{or} \qquad E = h\nu = \frac{hc}{\lambda}$$

$$\text{But,} \quad E = mc^2 = \frac{hc}{\lambda} \qquad \text{or} \quad mc = \frac{h}{\lambda}$$

If a photon of mass 'm' moving with a velocity 'c', then its momentum p = mc

$$\therefore p = \frac{h}{\lambda} \quad \text{or} \quad \lambda = \frac{h}{p}$$

But, if photon of mass 'm' moving with a velocity 'v', then its momentum is p = mv

$$\therefore p = \frac{h}{\lambda} \quad \text{or} \quad \lambda = \frac{h}{p} = \frac{h}{mv}$$

Where 'h' is a universal constant, called the Planck's constant. Its value is 6.625 x 10^{-34} Js, c is the speed of light.

Matter Waves

Dual Nature of Matter:

The wave theory of electromagnetic radiation satisfactorily explains the phenomena of reflection, refraction, interference, diffraction and polarization. But it failed to explain the phenomena of Photoelectric Effect, Compton Effect.

On the other hand, they were explained on the basis of quantum theory of radiation. According to which a beam of light of frequency ν consists of small packets each having energy hν called photon or quanta.

Sometimes these photons behave like a waves and sometimes like a corpuscles i.e., particles. Thus radiation have dual nature i.e., wave and particle or quantum nature.

Matter waves and their characteristics properties

In 1924 Louis de Broglie suggested that the particles like protons, electrons, & neutrons in motion exhibit characteristic properties of waves. Thus a moving particle can be associated with a wave or a wave can guide the motion of the particle. Hence the waves associated with the moving particles are known as de-Broglie waves or matter waves.

According to de-Broglie hypothesis, a particle of mass 'm' moving with velocity 'v' is associated with the wave. This wave is called matter wave. The wavelength of matter wave in terms of its momentum 'p' is,

$$\lambda = \frac{h}{mv} = \frac{h}{p}$$

1. Matter waves are the waves associated with moving particles.
2. Lighter the particles, greater is the wavelength associated with it, because here
$$\lambda \propto 1/m$$
3. Greater the velocity of the particle, smaller is the wavelength associated with the particle.
$$\lambda \propto 1/\text{v}$$
4. Matter waves are not electromagnetic waves. Since they don't depends on the charge of the particle.

5. The velocity of the matter waves is not constant. But it depends on the velocity of the particle.
6. Light wave has got same velocity, for all wavelengths. But in case of matter waves, the velocity is inversely proportional to the wavelength.
7. It is not possible to determine the exact position and momentum of a moving particle simultaneously.
8. Matter waves are also called as de-Broglie waves (or) pilot waves.

Note: (i) : If a particle of mass 'm' moving with a velocity 'v', then its kinetic energy

$$E = \tfrac{1}{2} mv^2 \quad \text{or} \quad m^2v^2 = 2mE \quad \text{or} \quad mv = \sqrt{2mE}$$

$\therefore \lambda = \dfrac{h}{mv} = \dfrac{h}{\sqrt{2mE}}$ is the expression for deBroglie wavelength in terms of kinetic energy E.

(ii) : we know that, if an electron is accelerated under the potential difference of V, then the energy acquired by it will be 'eV', then

$$\therefore eV = \frac{1}{2}mv^2 = \frac{1}{2}\frac{m^2v^2}{m} = \frac{p^2}{2m}$$

$$or \quad p^2 = 2meV \quad \text{or} \quad p = \sqrt{2meV}$$

$$\therefore \lambda = \frac{h}{p} = \frac{h}{\sqrt{2meV}}$$

This is the expression for deBroglie wavelength in terms of accelerating potential V.

Problems:-

(1) Calculate the de-Broglie wavelength associated with an electron having a kinetic energy of 100 eV.

Data:- $\lambda = ?$, E=100eV=100 x 1.602 x10^{-19}J

$$\lambda = \frac{h}{\sqrt{2mE}} = \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times100\times1.602\times10^{-19}}} = 1.228\times10^{-10}m = 1.228\,\text{A}^\circ$$

(2) (May22) Evaluate the de-Broglie wavelength of Helium Nucleus accelerated through a potential difference of 500 V.

Data:- $\lambda = ?$, V=500V, m=$4m_p$=4x1.673x10^{-27}kg

$$\lambda = \frac{h}{\sqrt{2meV}} = \frac{6.63\times10^{-34}}{\sqrt{2\times4\times1.673\times10^{-27}\times1.602\times10^{-19}\times500}} = 6.403\times10^{-13}m = 6403\text{A}^\circ$$

(3) Calculate the de-Broglie wavelength of an electron accelerated under a potential difference of 100 V.

Soln.: $\lambda = ?$, V=100V

$$\lambda = \frac{h}{\sqrt{2meV}} = \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times1.602\times10^{-19}\times100}} = 1.228\times10^{-10}m = 1.228\text{A}^\circ$$

(4) Compute the de-Broglie wavelength for a neutron moving with one tenth part of the velocity of light. Given mass of neutron=1.674x10^{-27}kg.

Soln.: $\lambda = ?$, $v = \frac{c}{10}$ *and* $m_n = 1.674\times10^{-27}\,kg$

$$\lambda = \frac{h}{mv} = \frac{6.63\times10^{-34}}{1.674\times10^{-27}\times\frac{3\times10^{8}}{10}} = 1.320\times10^{-14}\,m$$

(5) Find the KE of a neutron which has a wavelength of 3Å. Given its mass=1.674x10^{-27}kg.

$$\lambda = \frac{h}{\sqrt{2mE}} \quad \Rightarrow \quad E = \frac{h^2}{2m\lambda^2}$$

$$= \frac{\left(6.63\times10^{-34}\right)^2}{2\times1.674\times10^{-27}\times\left(3\times10^{-10}\right)^2} = 1.46\times10^{-21}J = 9.114\times10^{-3}eV$$

(6) Estimate the potential difference through which a proton is needed to be accelerated so that its de- Broglie wavelength becomes equal to 1 Å, given that its mass is1.673x10^{-27}kg.

$$\lambda = \frac{h}{\sqrt{2meV}} \quad \Rightarrow \quad V = \frac{h^2}{2me\lambda^2}$$

$$= \frac{\left(6.63\times10^{-34}\right)^2}{2\times1.673\times10^{-27}\times1.602\times10^{-19}\times\left(1\times10^{-10}\right)^2} = 0.082V.$$

(7) The kinetic energy of an electron is equal to the energy of photon with a wavelength of 560nm. Calculate the de-Broglie wavelength of the electron.

$$E = h\nu = \frac{hc}{\lambda}, \lambda_p = 560nm = 560\times10^{-9}\,m, \lambda = ?$$

$$E = \frac{hc}{\lambda_p} = \frac{6.63\times10^{-34}\times3\times10^{8}}{560\times10^{-9}} = 3.552\times10^{-19}J = 2.217eV$$

$$\lambda = \frac{h}{\sqrt{2mE}} = \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times3.552\times10^{-19}}} = 8.25\times10^{-10}\,m$$

(8) Calculate the de-Broglie wavelength associated with an electron having a kinetic energy of 100 eV.

Soln.: $\lambda = ?$, E=100eV=100 x 1.602 x10^{-19}J

$$\lambda = \frac{h}{\sqrt{2mE}}$$

$$\lambda = \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times100\times1.602\times10^{-19}}}$$

$$\lambda = 1.228\times10^{-10}\,m$$

(9) (May22) Evaluate the de-Broglie wavelength of Helium Nucleus accelerated through a potential difference of 500 V.

Soln.: $\lambda = ?$, V = 500V, m = $4m_p = 4x1.673x10^{-27}$kg

$$\lambda = \frac{h}{\sqrt{2meV}}$$
$$= \frac{6.63\times10^{-34}}{\sqrt{2\times4\times1.673\times10^{-27}\times1.602\times10^{-19}\times500}}$$
$$= 6.403\times10^{-13}\,m$$

(10) Calculate the de-Broglie wavelength of an electron accelerated under a potential difference of 100 V.

Soln.: $\lambda = ?$, V=100V

$$\lambda = \frac{h}{\sqrt{2meV}}$$
$$= \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times1.602\times10^{-19}\times100}}$$
$$= 1.228\times10^{-10}\,m$$

(11) Compute the de-Broglie wavelength for a neutron moving with one tenth part of the velocity of light. Given mass of neutron=1.674X 10^{-27}kg.

Soln.: $\lambda = ?$, $v = \frac{c}{10}$ *and* $m_n = 1.674\times10^{-27}\,kg$

$$\lambda = \frac{h}{mv}$$
$$= \frac{6.63\times10^{-34}}{1.674\times10^{-27}\times\frac{3\times10^{8}}{10}}$$
$$= 1.320\times10^{-14}\,m$$

(12) Find the KE of a neutron which has a wavelength of 3Å. Given its mass=1.674X 10^{-27}kg.

$$\lambda = \frac{h}{\sqrt{2mE}} \Rightarrow$$
$$E = \frac{h^2}{2m\lambda^2}$$
$$= \frac{\left(6.63\times10^{-34}\right)^2}{2\times1.674\times10^{-27}\times\left(3\times10^{-10}\right)^2}$$
$$= 1.46\times10^{-21}\,J$$

(13) Estimate the potential difference through which a proton is needed to be accelerated so that its de- Broglie wavelength becomes equal to 1Å, given that its mass is1.673X 10^{-27} kg.

$$\lambda = \frac{h}{\sqrt{2meV}} \Rightarrow$$

$$V = \frac{h^2}{2me\lambda^2}$$

$$= \frac{\left(6.63\times10^{-34}\right)^2}{2\times1.673\times10^{-27}\times1.602\times10^{-19}\times\left(1\times10^{-10}\right)^2}$$

$$= 0.082V.$$

(14) The kinetic energy of an electron is equal to the energy of photon with a wavelength of 560nm. Calculate the de-Broglie wavelength of the electron.

$$E = h\nu = \frac{hc}{\lambda}, \lambda_p = 560nm = 560\times10^{-9}m, \lambda = ?$$

$$E = \frac{hc}{\lambda_p}$$

$$E = \frac{6.63\times10^{-34}\times3\times10^{8}}{560\times10^{-9}}$$

$$E = 3.552\times10^{-19}J$$

$$\lambda = \frac{h}{\sqrt{2mE}}$$

$$\lambda = \frac{6.63\times10^{-34}}{\sqrt{2\times9.11\times10^{-31}\times3.552\times10^{-19}}}$$

$$\lambda = 8.25\times10^{-10}m$$

Wave Packet: A wave packet consisting of waves of slightly differing wavelengths may represent the moving particle. Superposition of these waves constituting the wave packet results in the net amplitude being modified, thereby defining the shape of the wave group.

A wave is represented by the formula

$$y = ASin(\omega t - kx)$$

Where y is the displacement at any instant t, A is the amplitude of vibration, ω is the angular frequency (ω = 2πν) and k is the wave vector (k = 2π/λ).

Phase Velocity (v_{Phase} or v_p)

A point marked on a wave can be regarded as representing a particular phase for the wave at that point. The velocity with which such a point would propagate is known as phase velocity (or) wave velocity. It is represented by

$$v_{phase} \quad \text{or} \quad v_p = \frac{\omega}{k}$$

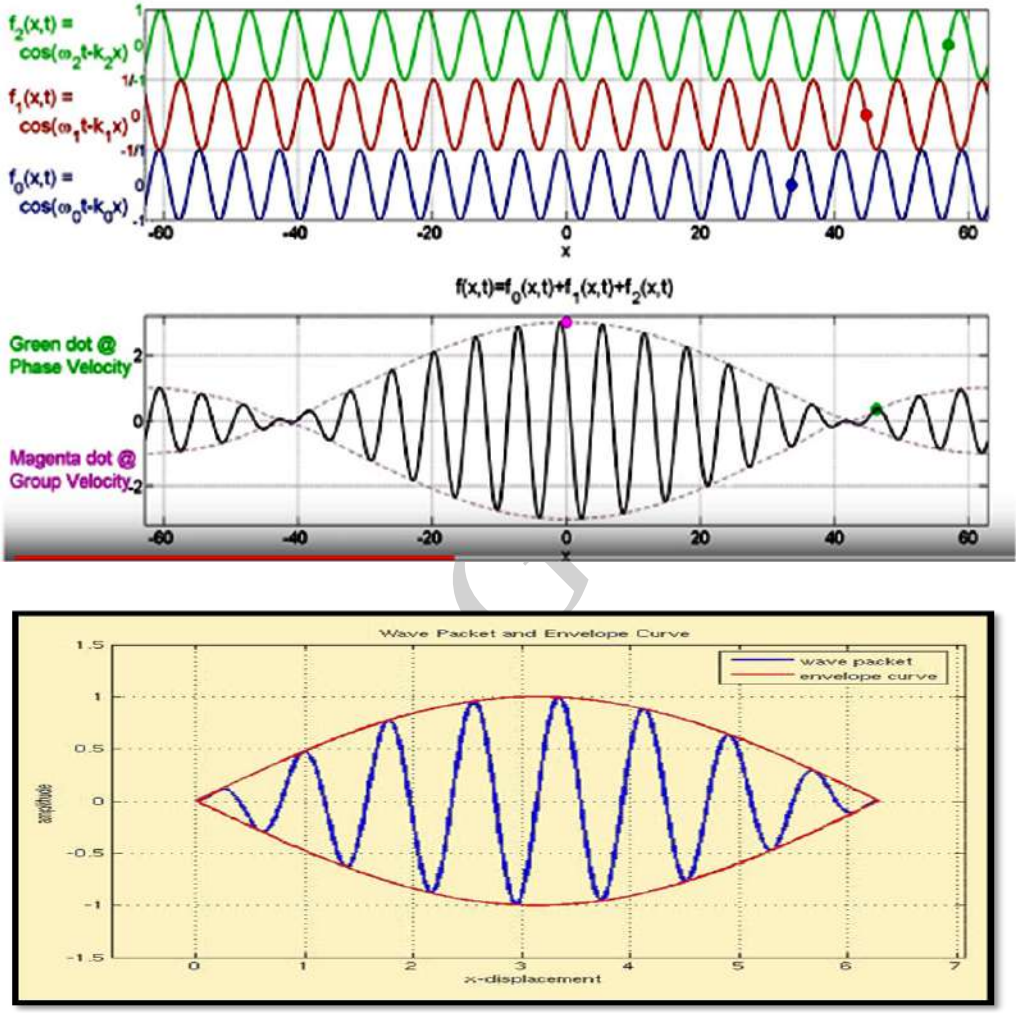
where, ω is angular frequency and k is the propagation constant or wave number

Group Velocity (v_{group} or v_g)

The velocity with which the resultant envelops of the group of waves travels is called group velocity.

It is denoted by v_g or v_{group} and is equal to the particle velocity v.

$$v_{group} \quad or \quad v_g = \frac{d\omega}{dk}$$



Relation between Group Velocity v_g and Phase Velocity v_p

The equations for group velocity and phase velocity are given by,

$$v_g = \frac{dw}{dk} \qquad \rightarrow (1)$$

$$v_p = \frac{\omega}{k} \qquad \rightarrow (2)$$

where ω is the angular frequency of the wave and k is the propagation constant or wave vector.

$$\therefore \omega = v_p k \qquad \rightarrow (3)$$

$$\frac{dw}{dk} = \frac{d\left(v_p k\right)}{dk} = v_p + k\frac{dv_p}{dk} \quad \rightarrow (4)$$

$$\text{But,} \quad k\frac{dv_p}{dk} = k\frac{dv_p}{d\lambda} \times \frac{d\lambda}{dk} \quad \rightarrow (5)$$

$$\text{we know that,} \quad k = \frac{2\pi}{\lambda} \quad or \quad \lambda = \frac{2\pi}{k}$$

$$\text{and} \quad \frac{dk}{d\lambda} = 2\pi\left(\frac{-1}{\lambda^2}\right) \quad or \quad \frac{d\lambda}{dk} = -\frac{\lambda^2}{2\pi}$$

$$\therefore \text{eqn (5) becomes,} \quad k\frac{dv_p}{dk} = \frac{2\pi}{\lambda} \times \frac{-\lambda^2}{2\pi} \times \frac{dv_p}{d\lambda}$$

$$or \quad k\frac{dv_p}{dk} = -\lambda \times \frac{dv_p}{d\lambda}$$

On substituting this values in equation (4) we get,

$$v_g = v_p - \lambda\frac{dV_p}{d\lambda}$$

$$or \quad v_{group} = v_{phase} - \lambda\left(\frac{dv_{phase}}{d\lambda}\right)$$

This is the relation between group velocity and phase velocity.

Heisenberg's uncertainty principle

According to this principle "It is impossible to determine precisely and simultaneously the values of both the members of the pair of physical variables, which describe the motion of the atomic system". Such variables are called canonically conjugate variables.

Example: Position and momentum, energy and time etc.,

Statement: *"it is impossible to determine simultaneously both position and momentum of a moving particle accurately at same time. The product of uncertainty in these quantities is always greater than or equal to h/4π".*

If Δx and ΔP_x are the uncertainties in the measurement of position and momentum of a particle, then

$$\Delta x . \Delta p_x \geq \frac{h}{4\pi}$$

If Δx is small, ΔP_x will be large and vice versa. That is if one quantity is measured accurately, the other quantity becomes less accurate.

Similarly the other uncertainty relations for other physical variables pair are,

$$\Delta E.\Delta t \geq \frac{h}{4\pi}$$

$$\Delta L.\Delta \theta \geq \frac{h}{4\pi}$$

Applications of Uncertainty Principle:

Non-existence of electrons in the nucleus and its implications non-relativistic approach

According to theory of relativity, if a particle of mass moving with a velocity v ,then the energy E and momentum p of the particle are expressed as,

$$E = mc^2 \quad \text{and} \quad p = mv \qquad \rightarrow (1)$$

where m is the relativistic mass of a particle moving the a velocity v and the expression for it in terms of rest mass m_0 can be written as,

$$m = \frac{m_o}{\sqrt{1-\frac{v^2}{c^2}}} \qquad \rightarrow (2)$$

$$m^2 = \frac{m_0^2}{1-\frac{v^2}{c^2}} = \frac{m_0^2 c^2}{c^2 - v^2}$$

$$m^2(c^2 - v^2) = m_0^2 c^2 \quad \text{or} \quad m^2c^2 - m^2v^2 = m_0^2c^2 \quad \text{x } c^2$$

$$m^2c^4 - m^2v^2c^2 = m_0^2c^4$$

Since $E = mc^2$ and $p = mv$, the above equation becomes

$$E^2 - p^2c^2 = m_0^2c^4 \quad \text{or} \quad E^2 = p^2c^2 + m_0^2c^4$$

$$E = c\sqrt{p^2 + m_0^2c^2} \quad \text{or} \quad E = c\left(p^2 + m_0^2c^2\right)^{1/2} \qquad \rightarrow (3)$$

According to Heisenberg's uncertainty principle we have,

$$\Delta x.\Delta p_x \geq \frac{h}{4\pi} \quad \text{or} \quad \Delta p_x \geq \frac{h}{4\pi.\Delta x} \qquad \rightarrow (4)$$

We know that the size / diameter of the nucleus is of the order of 10^{-14} m. If an electron is to exist inside the nucleus, then the uncertainty in its position Δx must not exceed the size of the nucleus,

$$i.e., \quad \Delta x \leq 10^{-14} m$$

Using Δx in equation (5) we have,

$$\Delta p_x \geq \frac{h}{4\pi.\Delta x} \geq \frac{6.63\times10^{-34}}{4\pi\times10^{-14}} \geq 0.5\times10^{-20} \text{ Ns}$$

$$\therefore \quad p_x \geq 0.5\times10^{-20} \ Ns$$

This is the uncertainty in momentum of an electron and it is equal to the momentum of the electron inside the nucleus,

Using momentum value P_x in an equation (3) we get,

$$E = 3x10^8\left((0.5x10^{-20})^2 + (9.11x10^{-31})^2 x(3x10^8)^2\right)^{1/2}$$

$$E = 1.5x10^{-12}\,J = \frac{1.5x10^{-12}}{1.602\,x10^{-19}}\,eV$$

$$E = 9.4x10^{6}\,eV \approx 9.4MeV$$

An electron may exist inside the nucleus if its energy is equal to or greater than 9.4 MeV. But, the experimental investigations on β-decay say that kinetic energy of the β–particles is 3 to 4 MeV. This clearly indicates that, electrons cannot exist within the nucleus.

$$n \rightarrow p + e^{-1} + \bar{\nu}\ (energy) \qquad \text{or} \qquad p \rightarrow n + e^{+1} + \nu\ (energy)$$

Principle of Complementarity

Statement: Bohr stated as "In a situation where the wave aspect of a system is revealed, its particle aspect is concealed; and in a situation where the particle aspect is revealed, its wave aspect is concealed. Revealing both simultaneously is impossible; the wave and particle aspects are complementary."

Explanation: We know that the consequence of the uncertainty principle is both the wave and particle nature of the matter cannot be measured simultaneously. In other words, we cannot precisely describe the dual nature of Light.

• If an experiment is designed to measure the particle nature of the matter, during this experiment, errors of measurement of both position and the time coordinates must be zero and hence the momentum, energy and the wave nature of the matter are completely unknown.

• Similarly, if an experiment is designed for measuring the wave nature of the particle, then the errors in the measurement of the energy and the momentum will be zero, whereas the position and the time coordinates of the matter will be completely unknown. unknown.

From the above explanation, we can conclude that, when the particle nature of the matter is measured or displayed, the wave nature of the matter is necessarily suppressed and vice versa.

Problems:

1. If the group velocity of a particle is $3x10^6$ m/s, calculate its phase velocity. (Given, c = $3x10^8$ m/s).

$$Given,\ v_g = 3\times10^6\,m/s,\ \ c = 3\times10^8\,m/s,\ \ v_p = ?$$

$$w.k.t; \qquad V_{group}\,V_{phase} = c^2$$

$$or \quad V_{phase} = \frac{c^2}{V_{group}} = \frac{(3\times10^8)^2}{(3\times10^6)} = \frac{9\times10^{16}}{3\times10^6} = 3\times10^{10}\,m/s$$

2. Calculate the de-Broglie wavelength associated with a proton moving with a velocity equal to 1/20 th of the velocity of light, if the mass of the proton is $1.67x10^{-27}$ kg.

$$Given,\ \ v = \frac{1}{20}\times c = \frac{1}{20}\times3\times10^8,\ \ m_p = 1.67\times10^{-27}\,kg,\ \ \lambda_p = ?$$

$$w.k.t,\ \ \lambda_p = \frac{h}{mv}$$

$$\lambda_p = \frac{6.63\times10^{-34}}{1.67\times10^{-27}\times\frac{1}{20}\times3\times10^8} = \frac{20\times6.63\times10^{-34}}{1.67\times10^{-27}\times3\times10^8} = 2.646\times10^{-14}\,m$$

3. Find the kinetic energy and group velocity of an electron with de-Broglie wavelength of 0.2 nm.

$Given\ ,\quad \lambda = 0.2\ nm = 0.2\times10^{-9}\,m,\quad K.E = ?,\quad V_{group} = ?,\quad m_e = 9.1\times10^{-31}\,kg$

$$w.k.t;\quad \lambda = \frac{h}{p}$$

$$\therefore\ p = \frac{h}{\lambda} = \frac{6.63\times10^{-34}}{0.2\times10^{-9}} = 3.315\times10^{-24}\,kg-m/s$$

$$K.E = E = \frac{p^2}{2m} = \frac{\left(3.315\times10^{-24}\right)^2}{2\times9.1\times10^{-31}} = 6.038\times10^{-18}\,J$$

$$V_g\ or\ V_{group} = \frac{p}{m} = \frac{3.315\times10^{-24}}{9.1\times10^{-31}} = 3642857.143 = 3.64\times10^{6}\,m/s$$

5. If the uncertainty in the position of an electron is 4×10^{-10} m, calculate the uncertainty in its momentum.

$Given,\quad \Delta x = 4\times10^{-10}m,\quad \Delta P_x = ?$

$$\text{w.}k.t\quad \Delta x.\Delta p_x \geq \frac{h}{4\pi}$$

$$\therefore \Delta p_x \geq \frac{h}{4\pi\Delta x} \geq \frac{6.63\times10^{-34}}{4\times3.14\times4\times10^{-10}} \geq 1.318\times10^{-25}\,kgs^{-1}$$

6. In a simultaneous measurement of position and velocity of an electron moving with a speed of 6×10^5 m/s. Calculate the highest accuracy with which its position could be determined if the inherent error in the measurement of the velocity is 0.01% for the speed stated.

$Given,\quad v = 6\times10^5 m/s,\quad \%\ of\ error\ in\ velocity = 0.01\%,\quad \Delta x = ?$

$$percentage\quad of\ uncerta\ \text{int}\ y\ \ in\ \ velocity\ = \frac{uncerta\ \text{int}\ y\ \ in\ \ velocity}{velocity}\times100$$

$$uncertaini\ \ ty\ \ in\ \ velocity\ = \frac{percentage\ \ of\ uncertaini\ \ ty\ \ in\ \ velocity}{100}\times velocity$$

$$\Delta v = \frac{\%\ of\ \ error\ \ in\ the\ measurment\ \ \ of\ \ velocity}{100}\times v$$

$$\Delta v = \frac{0.01}{100}\times6\times10^5 = 60\,ms^{-1}$$

$$\text{w.}k.t;\quad \Delta x.\Delta p_x \geq \frac{h}{4\pi}$$

$$\Delta x \geq \frac{h}{4\pi.m.\Delta v_x} \geq \frac{6.63\times10^{-34}}{4\times3.142\times9.1\times10^{-31}\times60} \geq 9.66\times10^{-7}\,m = 966\ nm$$

7. An electron has a speed of 300 m/s accurate to 0.01% with what fundamental accuracy can we locate the position of the electron.

Given, V = 300 m/s, Δx = ?, Δv = ? and % of accuracy in speed = 0.01

$$\Delta v = 0.01\%\ of\ v \quad \text{or} \quad \Delta v = \frac{0.01}{100} \times v = \frac{0.01}{100} \times 300 = 0.03\ m/s$$

$$\Delta v = 3 \times 10^{-2}\ m/s$$

w.k.t

$$\Delta x.\Delta p_x \geq \frac{h}{4\pi} \quad \text{or} \quad \Delta x \geq \frac{h}{4\pi\,\Delta p_x} = \frac{h}{4\pi.m.\Delta v}$$

$$\therefore \Delta x \geq \frac{6.63 \times 10^{-34}}{4 \times 3.142 \times 9.1 \times 10^{-31} \times 3 \times 10^{-2}} = 1.932 \times 10^{-3}\ m$$

$\therefore$ The maximum accuracy with which the electron can be located is 1.932 x 10^{-3} m.

8. The speed of electron is measured to within an uncertainty of 1x10^4 m/s. What is the minimum space required by the electron to be confined in an atom?

Given, $\Delta v = 1 \times 10^4\ m/s, \Delta x = ?$

w.k.t; $\Delta x.\Delta p_x \geq \frac{h}{4\pi}$

$$\therefore\ \Delta x \geq \frac{h}{4\pi\,\Delta p_x} = \frac{h}{4\pi.m.\Delta v} \geq \frac{6.63 \times 10^{-34}}{4 \times 3.142 \times 9.1 \times 10^{-31} \times 1 \times 10^4} = 5.797 \times 10^{-9}\ m$$

$$\Delta x \geq 57.97 \times 10^{-10}\ m = 57.97\ \overset{0}{A}$$

9. The position and momentum of 1keV electron are simultaneously determined and if its position is located within 1Å. What is the minimum percentage of uncertainty in its momentum?

Given, E = 1keV, E = 1x10^3 eV, E = 1x10^3 x1.602x10^{-19} J, E = 1.602x 10^{-16} J

$\Delta x = 1\ \overset{0}{A} = 1 \times 10^{-10}\ m \qquad \Delta P = ?\ , \frac{\Delta P}{P} \times 100 = ?$

$$w.k.t; \quad \Delta x.\Delta p_x \geq \frac{h}{4\pi}$$

$$\therefore \quad \Delta P_x \geq \frac{h}{4\pi\,\Delta x} \geq \frac{6.63 \times 10^{-34}}{4 \times 3.142 \times 10^{-10}} = 5.275 \times 10^{-25}\ \ kgms^{-1}$$

we have the equation for momentum,

$$P = \sqrt{2mE} = \sqrt{2 \times 9.1 \times 10^{-31} \times 1.602 \times 10^{-16}}$$

$$P = 1.707 \times 10^{-23}\ kg\ ms^{-1}$$

$$Percentage\ of\ uncertainty\ in\ momentum = \frac{Uncertainty\ in\ momentum}{momentum} \times 100$$

$$Percentage\ of\ uncertainty\ in\ momentum = \frac{5.275 \times 10^{-25}}{1.707 \times 10^{-23}} \times 100 = 3.087 = 3.1$$

10. The inherent uncertainty in the measurement of time spent by Iridium -191 nuclei in the excited state is found to be 1.4x10^{-10}s. Estimate the uncertainty that results in its energy in the excited state.

$Given,\ \ \Delta t = 1.4\times10^{-10}\,s,\ \ \Delta E = ?$

$$\text{w.k.t;}\quad \Delta E.\Delta t \geq \frac{h}{4\pi}$$

$$\therefore\quad \Delta E \geq \frac{h}{4\pi.\Delta t} \geq \frac{6.63\times10^{-34}}{4\times3.142\times1.4\times10^{-10}} = 3.768\times10^{-25}\,J$$

$$\Delta E = \frac{3.768\times10^{-25}}{1.602\times10^{-19}}\,eV = 2.353\times10^{-6}\,eV$$

11. The average time that an atom retains excess excitation energy before re- emitting it in the form of electromagnetic radiation is 10^{-8}s. Calculate the limit of accuracy with which the excitation energy of the emitted radiation can be determined.

(OR)

What is the minimum uncertainty in the energy state of an atom if an electron remains in this state for 10^{-8} seconds?

$Given,\ \ \Delta t = 1.4\times10^{-10}\,s,\ \ \Delta E = ?$

$$\text{w.k.t;}\quad \Delta E.\Delta t \geq \frac{h}{4\pi}$$

$$\therefore\quad \Delta E \geq \frac{h}{4\pi.\Delta t} \geq \frac{6.63\times10^{-34}}{4\times3.142\times10^{-8}} = 5.275\times10^{-27}\,J$$

$$\Delta E = \frac{5.275\times10^{-27}}{1.602\times10^{-19}}\,eV = 3.292\times10^{-8}\,eV$$

12. An electron is confined to a box of length 10^{-8}m. Calculate the minimum uncertainty in its velocity.

$Given,\ \ \Delta v = ?\ \ \Delta x = 10^{-8}\,m$

$$w.k.t;\ \ \Delta x.\Delta p_x \geq \frac{h}{4\pi}$$

$$or\ \ \Delta P_x \geq \frac{h}{4\pi\,\Delta x}\ \text{or}\quad m.\Delta v_x = \frac{h}{4\pi.\Delta x}$$

$$\therefore\ \ \Delta v_x = \frac{h}{4\pi.m.\Delta x} \geq \frac{6.63\times10^{-34}}{4\times3.142\times10^{-8}\times9.1\times10^{-31}} = 5797.03\ \ m/s$$

$$\Delta v = 5800\ \ m/s$$

(13) If the kinetic energy of an electron known to be about 1eV, must be measured to within 0.0001eV, What accuracy can its position be measured simultaneously?

Soln.: E = 1eV = $1.602\text{x}10^{-19}$J, ΔE = 0.0001eV = $0.0001\text{x}1.602\text{x}10^{-19}$J, Δx = ?

$$\Delta P\Delta X = \frac{h}{4\pi}\quad and\quad P = \sqrt{2mE} \Rightarrow$$

$$\Delta P = \frac{1}{2}(2mE)^{-1/2}\,2m\Delta E$$

$$= \sqrt{\frac{m}{2E}}\,\Delta E$$

$$\therefore \Delta X = \frac{h}{4\pi}\frac{1}{\Delta E}\sqrt{\frac{2E}{m}}$$

$$\therefore \Delta X = \frac{6.63\times10^{-34}\times\sqrt{2\times1.602\times10^{-19}}}{4\times\pi\times0.0001\times1.602\times10^{-19}\times\sqrt{9.1\times10^{-31}}}$$

$$= 1.95\times10^{-6}\,m$$

(14) Uncertainty in time of an exited atom is about 10^{-8}s. What are the uncertainties in energy and in frequency of the radiation?

Soln.: $\Delta t = 10^{-8}$s, ΔE = ?, $\Delta\nu$ = ?

$$\Delta E\Delta t = \frac{h}{4\pi}$$

$$\Delta E = \frac{h}{4\pi.\Delta t}$$

$$= \frac{6.63\times10^{-34}}{4\times\pi\times10^{-8}}$$

$$= 5.23\times10^{-27}\,J$$

$$E = h\nu$$

$$\therefore \Delta E = h\Delta\nu$$

$$\therefore \Delta\nu = \frac{\Delta E}{h}$$

$$= \frac{5.23\times10^{-27}}{6.63\times10^{-34}}$$

$$= 7.89\times10^{6}\,Hz$$

(3) (May22) The position and momentum of an electron with energy 0.5 keV is found with a minimum percentage uncertainty in momentum. Find its uncertainty if the measurement of position has an uncertainty of 0.5Å.

Soln.: E=0.5keV=0.5x10^{3}x1.602x10^{-19}J, Percentage uncertainty in momentum = ?,

Δx =0.5x10^{-10} m.

$$\Delta P\Delta X = \frac{h}{4\pi} \quad and \quad P = \sqrt{2mE}$$

$$\Delta P = \frac{h}{4\pi\times\Delta X} \quad and \quad P = \sqrt{2mE}$$

$$\Delta P = \frac{6.63\times10^{-34}}{4\pi\times0.5\times10^{-10}} \quad and \quad P = \sqrt{2\times9.1\times10^{-31}\times0.5\times10^{3}\times1.602\times10^{-19}}$$

$$\Delta P = 1.055\times10^{-24} \quad and \quad P = 1.207\times10^{-23}$$

Questions

Short Answer Questions

1. What is Planck's law of radiation?
2. List out the characteristics of matter waves.
3. State and explain Planck's law of radiation.
4. What are matter waves?
5. Define phase velocity and group velocity.
6. Obtain an expression for deBroglie wavelength.
7. State and Explain deBroglie hypothesis.
8. What are matter waves and mention their properties?
9. Derive the relation between Phase velocity and Group velocity.
10. What is de-Broglie concept of matter wave? Explain the characteristics of matter wave.
11. Define phase velocity and group velocity & obtain a relation between them.
12. State Heisenberg Uncertainty Principle.
13. Show that the electron does not exists inside the nucleus of an atom.
14. Explain the principle of complementarity.
15. State and Explain Heisenberg's uncertainty principle.
16. Explain why electron cannot exist inside the nucleus?
17. State and Explain the Principle of Complementarity.

Problems:

1. Compare the energy of photon with that of an electron when both are associated with a wavelength of 0.2 nm.

2. Calculate the deBroglie wavelength of a 1000 kg automobile travelling at 100 m/s and a 0.1 kg bullet travelling at 500 m/s.

3. A fast moving neutron is found to a have an associated deBroglie wavelength of 2×10^{-12} m. Find its kinetic energy and group velocity of the deBroglie waves using the relativistic change in mass. (Mass of neutron = 1.675×10^{-27} kg)

6. Calculate the deBroglie wavelength associated with an electron with a kinetic energy of 2000 eV.

7. Calculate the momentum and the deBroglie wavelength of the particle associated with an electron with a kinetic energy of 1.5 keV.

8. Calculate the wavelength associated with an electron having kinetic energy 100 eV.

9. Calculate deBroglie wavelength associated with electron carrying energy 2000 eV.

10. Find the energy of the neutron in eV whose deBrogle wavelength is 1 Å.

11. Calculate deBroglie wavelength associated with neutron of mass 1.674×10^{-27} kg with $1/10^{th}$ part of speed of light.

12. Calculate deBroglie wavelength associated with electrons whose speed is 0.01 part of the speed of light.

13. What is the deBroglie wavelength of a proton whose energy is 3eV given mass of proton is 1.67×10^{-27} kg.

14. Find the kinetic energy and group velocity of an electron with deBroglie wavelength of 0.2 nm.

15. Calculate the deBroglie wavelength of particle of mass 0.65 MeV/c^2 has a kinetic energy 80 eV.

16. Find deBroglie wavelength of a particle of mass 0.58 MeV/c^2 has a kinetic energy 90 eV, where c is speed of light.

17. A particle of mass 940 MeV/c^2 has kinetic energy 0.5 keV. Find it deBroglie wavelength, where c is speed of light.

18. Find the deBroglie wavelength of an electron accelerated through a potential difference of 182 V and object of mass 1 kg moving with a speed of 1 m/s. Compare the results and comment.

19. The position and momentum of an electron with energy 0.5 keV are determined. What is the minimum percentage uncertainty in its momentum if the uncertainty in the measurement of position is 0.5Å?

20. The speed of electron is measured to within an uncertainty of 2.2×10^4 m/s in one dimension. What is the minimum width required by the electron to be confined in an atom?

Wave Function

Introduction:

In general, a wave is characterised by periodic variation in some physical quantity.

For example – pressure varies periodically in sound waves whereas electric and magnetic fields vary periodically in an electromagnetic wave. Similarly, whose periodic variations make up the matter wave is called wave function.

Wave Function:

The variable quantity that characterises the deBroglie wave is called wave function. Wave function in quantum mechanics accounts for the wave like properties of particle and is obtained by solving a fundamental equation called Schrödinger's equation.

The wave functions vary with respect to both position co-ordinates of the physical system and the time (x, y, z & t) is called total wave function.

It is denoted by the capital form of Greek letter 'Ψ'. If the wave function has variation only with position (x, y, z) it is denoted by the lower case Greek letter 'ψ'.

The total wave function can be mathematically represented by the equation

$$\Psi = Ae^{i(kx-\omega t)}$$

Where A is a constant and *w* is the angular frequency of the wave

The above equation can be written as

$$\Psi = Ae^{ikx}e^{-i\omega t}$$

Where, $\Psi = Ae^{ikx}$ is the space dependent wave function and is therefore time independent wave function. $e^{-i\omega t}$ is the time dependent wave function.

The wave function ψ is a measure of finding the particle at a particular position (x, y, z) and at time t.

The following are the basic properties of wave function:

1. The wave function ψ itself does not have any physical significance.
2. ψ is a positive or negative or complex quantity and hence it cannot be measured.
3. ψ is a function of space and time (r, t) coordinates and describes the behaviour of a single particle or photon and wave nature.
4. ψ is a large magnitude where the particle (Photon, electron etc.,) to be located and small at other places.
5. The probability of finding a particle at some point in space at time 't' is a positive value between 0 & 1; i.e., $|\psi|^2$ is real and +ve between 0 & 1.

Probability Density:

In classical mechanics, the square of wave amplitude associated with electromagnetic radiation is interpreted as measure of intensity. This suggests there will be a similar interpretation for de-Broglie waves associated with electron or any particle.

Let τ be a volume inside which a particle is present, but where exactly the particle is situated inside τ is not known

"If ψ is the wave function associated with the particle then the probability of finding the particle in certain volume dτ of τ is equal to $|\psi|^2$ dτ. So $|\psi|^2$ is called the probability density".

$$|\psi|^2 \, d\tau$$

This interpretation was first given by Max Born in 1926.

If the value of $|\psi|^2$ is large at a point in a given time, then the probability of finding the particle at that point and time is more. If $|\psi|^2 = 0$, then the probability of finding the particle is zero or less.

Therefore the total wave function can be represented by the equation,

$$\Psi = Ae^{i(kx-\omega t)} \rightarrow (1)$$

where A is a constant, ω is angular frequency of the wave
The complex conjugate of Ψ is given by,

$$\Psi^* = Ae^{-i(kx-\omega t)} \rightarrow (2)$$

From equation (1) and (2), $\Psi\Psi^*$ is real and positive quantity which is called the probability density.

$$\text{i.e,} \quad |\psi|^2 = \psi\psi^* = A^2$$

Therefore $|\Psi|^2 dx$ is the probability density in 1- dimension,

and $|\Psi|^2$ dv is the probability density in 3-dimension.

Normalization:

According to Born's interpretation the probability of finding the particle within an element of volume is $|\Psi|^2$ dv, since the particle is certainly present somewhere inside the volume dv.

Therefore "The integral of the square of the wave function over the entire volume in space must be equal to unity" and mathematically it is represented as,

$$\int_{-\infty}^{\infty} |\psi|^2 \, d\text{v} = 1$$

Where, the wave function satisfying the above relation is the normalized wave function.

Very often Ψ is not a normalized wave function. If this function Ψ is multiplied by a constant A, then the new wave function $A\Psi$ is also a solution of the wave equation. Hence the new wave function is a normalized wave function, if

$$\int_{-\infty}^{\infty} |A\psi^*||A\psi| d\text{v} = 1 \quad \text{or} \quad A^2 \int_{-\infty}^{\infty} \psi\psi^* d\text{v} = 1$$

$$|A|^2 = \frac{1}{\int_{-\infty}^{\infty} \psi^* \psi \, dv}$$

Where $|A|^2$ is known as normalizing constant, the quantity $|A\Psi|^2$ represents probability.

Therefore, the process of constructing $A\Psi$ from Ψ is called normalization of the wave function.

Limitations of wave function:

1. The wave function Ψ must be finite for all values of x, y, z and it is finite for a particular point.
2. Ψ must be single valued everywhere for each set of x, y, z and must have unique value.
3. Ψ must be continuous in all regions except where the potential energy V is infinite.
4. Ψ and its first derivatives $d\Psi/dx$, $d\Psi/dy$, $d\Psi/dz$ must be continuous and single valued everywhere.
5. Ψ must be normalised and in order that $|\Psi|^2 dv$ over all space be a finite constant.

Expectation Value

In quantum mechanics, the expectation value is the probabilistic expected value of the result (measurement) of an experiment. It can be thought of as an average of all the possible outcomes of a measurement as weighted by their likelihood. Expectation value as such it is not the most probable value of a measurement. In the real sense the expectation value may have zero probability of occurring.

Let us consider a particle moving along the x axis. The result of a measurement of the position x is a continuous random variable. Consider a wave function $\Psi(x, t)$. The $|\Psi(x, t)|^2$ value is a probability density for the position observable and $|\Psi(x, t)|^2 dx$ is the probability of finding the particle between x and $x+dx$ at time t. Thus, if a measurement of position is repeated many times in an identical way on an identical particle in identical circumstances, many possible outcomes are possible and the expectation value of these outcomes is, according to the following equation

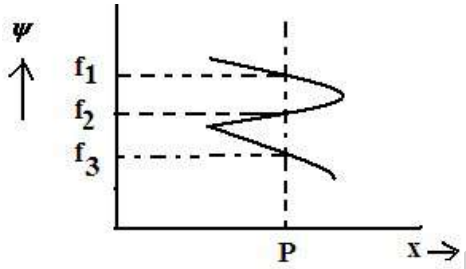
$$\langle x \rangle = \int_{-\infty}^{\infty} x |\psi(x, t)|^2 dx$$

Properties of Wave Functions:

A system is characterised by its position, energy, momentum etc,. In quantum mechanics, the state of a system is completely characterised by a wave function.

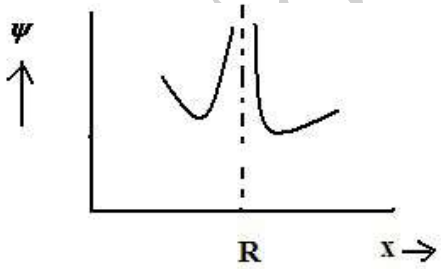
Physically acceptable wave function Ψ must satisfy the following conditions,

1. Ψ is single valued everywhere



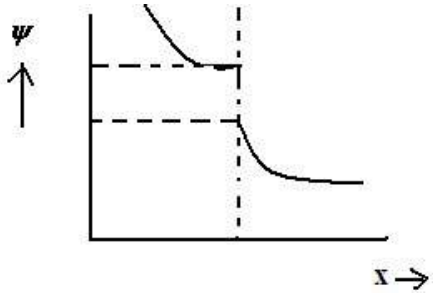
If Ψ has more than one value at any point (at P, Ψ has f_1, f_2 and f_3) it would mean more than one value of probability of finding the particle at that point which is obviously ridiculous. Therefore, Ψ must be single valued everywhere.

2. Ψ is finite everywhere



If Ψ is infinite at a point R there will be large probability of finding the particle at that point. This violates the uncertainty principle, therefore Ψ must have a finite or zero value at that point.

3. Ψ and its first derivatives dΨ/dx with respect to its variables are continuous everywhere



Ψ and its first derivatives must be continuous. Since the probability can have any value between zero and one, the wave function must be continuous.

Schrodinger's Time Independent One Dimensional Wave Equation

Based on de-Broglie idea of matter waves, Schrödinger developed a mathematical theory for a particle of mass 'm' moving with a velocity 'v' along x-direction associated with a wave of wavelength,

$$\lambda = \frac{h}{mv} = \frac{h}{p}$$

Where, p = mv is the momentum of the particle.

Let a wave function Ψ describing the de-Broglie wave travelling in +ve x-direction is given by,

$$\psi = Ae^{i(kx-\omega t)} \rightarrow (1)$$

Where Ψ is a total wave function, A is a constant and ω is angular frequency of wave.

Let us differentiate Ψ (in equation 1) twice with respect to 'x' then

$$\frac{d\psi}{dx} = A(ik)e^{i(kx-\omega t)}$$

$$\frac{d^2\psi}{dx^2} = A(ik)^2 e^{i(kx-\omega t)}$$

$$\frac{d^2\psi}{dx^2} = -k^2\psi \quad or \quad \frac{d^2\psi}{dx^2} + k^2\psi = 0 \rightarrow (2) \quad \because i^2 = -1$$

$$But \quad k = \frac{2\pi}{\lambda} \quad and \quad \lambda = \frac{h}{mv}$$

$$\therefore k = \frac{2\pi mv}{h} \quad \text{or} \quad k^2 = \frac{4\pi^2 m^2 v^2}{h^2}$$

Hence equation (2) becomes,

$$\frac{d^2\psi}{dx^2} + \frac{4\pi^2 m^2 v^2}{h^2}\psi = 0 \rightarrow (3)$$

The total energy E of the particle is the sum of kinetic energy T and potential energy V,

$$\therefore E = T + V$$

$$\text{But} \quad T = \frac{1}{2}mv^2 \quad \therefore \frac{1}{2}mv^2 = (E-V)$$

$$\text{or} \quad mv^2 = 2(E-V) \rightarrow (4)$$

Substitute this value of mv^2 in equation (3) we get

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E-V)\psi = 0 \rightarrow (5)$$

This is known as time independent 1 - dimensional Schrödinger equation.

Equation (5) can also be extended for 3-dimensional space as,

$$\frac{d^2\psi}{dx^2} + \frac{d^2\psi}{dy^2} + \frac{d^2\psi}{dz^2} + \frac{8\pi^2 m}{h^2}(E-V)\psi = 0 \rightarrow (6)$$

$$or \quad \nabla^2\psi + \frac{8\pi^2 m}{h^2}(E-V)\psi = 0 \rightarrow (7)$$

$$where \quad \nabla^2 = \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2}$$

Equation (6) and (7) are the 3-dimensional time independent Schrödinger wave equation, where Ψ is Ψ(x, y, z).

Eigen Functions and Eigen Values

"Eigen functions are those wave functions of quantum mechanics which possess the properties that they are single valued, finite everywhere and also their first derivatives with respect to their variables are continuous everywhere".

When the Eigen functions are operated by quantum mechanical operators on physical quantities like momentum, energy etc., of a system, the possible values are observed and these values are called Eigen values".

Ex: 1. If an operator say *d/dx* operates on a wave function $\Psi = e^{ax}$, then

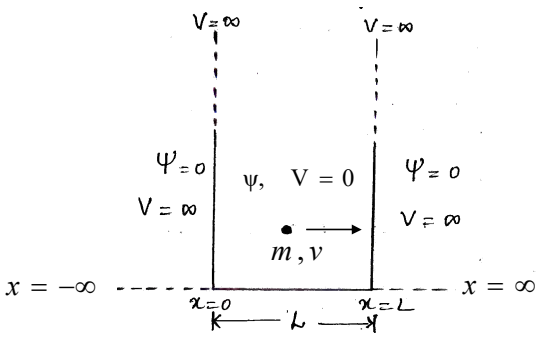
$$\frac{de^{ax}}{dx} = ae^{ax} = a\Psi$$

That is it produces the wave function multiplied by a constant. Such values obtained for a physical observable are called Eigen values.

Here 'a' is the Eigen value & $\Psi = e^{ax}$ is the Eigen function.

Applications of Schrödinger wave equation to particle trapped in a one dimensional square potential well

*** (Derivation of energy Eigen values and Eigen functions) ***



Consider a particle of mass 'm' moving with a speed 'v' along x-axis is confined to a box of length 'L' with perfectly rigid walls at $x = 0$ & $x = L$ as shown in the figure.

The particle does not lose energy when it collides with the walls so that its total energy remains constant. The potential energy V of the particle is constant within the box which can be taken to be zero for convenience.

$$\therefore V = 0 \quad for \quad 0 < x < L \quad \rightarrow (1)$$

The potential energy of the particle is infinite on and beyond the walls of the box.

$$\therefore V = \infty \quad for \quad x \le 0 \;\; and \;\; x \ge L \quad \rightarrow (2)$$

As the particle does not exist on the walls and beyond them, the wave function Ψ is zero.

$$\therefore \psi = 0 \quad for \quad x \le 0 \;\; x \ge L \quad \rightarrow (3)$$

The wave function Ψ exists within the box only.

$\therefore$ The Schrödinger's time independent wave equation is,

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0 \quad \rightarrow (4)$$

For the particle exists inside the box, V = 0

$\therefore$ Equation (4) becomes

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2} E\psi = 0 \quad \rightarrow (5)$$

$$\text{let} \quad \frac{8\pi^2 mE}{h^2} = k^2 \quad \rightarrow (6)$$

$\therefore$ Equation (5) becomes

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0 \quad \rightarrow (7)$$

This is the second order differential equation. The general solution of this equation is given by,

$$\psi = A\sin kx + B\cos kx \quad \rightarrow (8)$$

where, A & B are arbitrary constants, which are to be evaluated by using boundary conditions.

From the first boundary conditions, $\Psi = 0$ at $x = 0$, $\therefore$ Equation (8) becomes,

$$0 = A\sin 0 + B\cos 0$$

$$\text{Since,} \;\; \sin 0 = 0 \;\; \& \;\; \cos 0 = 1, \;\; \text{we have} \;\; B = 0$$

$\therefore$ Equation (8) becomes,

$$\psi = A \sin k x \quad \rightarrow (9)$$

From second boundary conditions, $\Psi = 0$ at $x = L$, $\therefore$ Equation (9) becomes,

$$0 = A \sin kL$$

OR

$$A \sin kL = 0$$

$$\because A \neq 0, \ \sin kl = 0 \quad \text{for all values of} \quad kL = n\pi \ \text{ where n} = 1, 2, 3, \ldots\ldots$$

$$\therefore k = \frac{n\pi}{L} \quad \rightarrow (10)$$

By substituting the value of k in equation (9) we get general wave function called Eigen wave function and Eigen energy equation.

$$\psi_n = A \sin\left(\frac{n\pi}{L}\right) x \rightarrow (11)$$

This is known as Eigen function or Eigen wave function.
Similarly by substituting the value of k in equation (6) we get,

$$\frac{8\pi^2 mE}{h^2} = \frac{n^2\pi^2}{L^2} \quad \text{or} \quad E = \frac{n^2 h^2}{8mL^2}$$

$$\text{In general} \quad E_n = \frac{n^2 h^2}{8mL^2} \quad \rightarrow (12)$$

This is the expression for Eigen values or Eigen energy values.

Thus, we see that in a potential well the particle cannot have an arbitrary energy, but it can have only discrete energy values corresponding to n = 1, 2, 3 ... are the Eigen values.

According to equation (11) if n = 0, $\Psi_n = 0$, which means that the particle doesn't present inside the box, which is not true. $\therefore$ The value of $E_n = 0$ for n = 0 is not acceptable. Hence the lowest allowed energy corresponding to n = 1 is called the 'zero-point energy or ground state energy'. Thus zero-point or ground state of energy of the particle in an infinite potential well is given by,

$$E_1 = \frac{h^2}{8mL^2}$$

The energy states corresponding to n >1 are called excited states.

Normalization:

To evaluate A in Eigen function Ψ_n, one has to perform the normalization of the wave function.

The allowed solutions of the Schrödinger equation are the Eigen functions, according to the equation.

$$\psi_n = A \sin\left(\frac{n\pi}{L}\right) x \quad \rightarrow (13)$$

The complex conjugate of equation (13) is,

$$\psi_n^* = A \sin\left(\frac{n\pi}{L}\right) x \quad \rightarrow (14)$$

To find the value of A, we use the normalization condition.

$$i.e., \quad \int_{-\infty}^{\infty} \psi_n^* \psi_n dx = \int_{-\infty}^{\infty} |\psi|^2 dx = 1$$

In this case, the particle exists only within the box of length (L). ∴The above equation can be written as,

$$\int_0^L |\psi_n|^2 dx = 1$$

By substituting the values of Ψ_n and Ψ_n^* in the above equation, we get

$$\int_0^L A^2 \sin^2 \frac{n\pi}{L} x \, dx = 1$$

$$A^2 \int_0^L \frac{1}{2}\left(1 - \cos\frac{2n\pi}{L}\right) x \, dx = 1 \qquad \left[\because \sin^2 A = \frac{1-\cos 2A}{2}\right]$$

$$\frac{A^2}{2}\left[\int_0^L dx - \int_0^L \cos\frac{2n\pi}{L} x dx\right] = 1$$

$$\frac{A^2}{2}\left[x - \frac{L}{2n\pi}\sin\left(\frac{2n\pi}{L}\right)x\right]_0^L = 1$$

$$\frac{A^2}{2}\left[L - \frac{L}{2n\pi}\sin(2n\pi) - 0 + 0\right] = 1 \quad \text{Here, for any value of n, } \sin 2n\pi = 0$$

$$\therefore \frac{A^2 L}{2} = 1 \quad \text{or} \quad A^2 = \frac{2}{L} \quad \text{or} \quad A = \sqrt{\frac{2}{L}}$$

Thus, by substituting the value of A in equation (13) we get normalized wave functions or Eigen function of a particle in one dimensional infinite potential well.

$$\psi_n = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}\right)x \quad \rightarrow (15)$$

The first three eigen functions Ψ_1, Ψ_2, Ψ_3 together with the probability densities $|\psi_1|^2$, $|\psi_2|^2$, $|\psi_3|^2$ and eigen values E_1, E_2, E_3 are as shown in figure (a), (b) & (c) respectively for n = 1, 2 & 3.

For n = 1, this is the ground state and the particle is normally found in this state.

$$\text{Eigen function} \quad \psi_1 = A\sin\left(\frac{\pi}{L}\right)x$$

$\Psi_1 = 0$ for both $x = 0$ and $x = L$ and Ψ_1 has maximum value A for $x = L/2$

At $x = 0$ and $x = L$, $|\Psi_1|^2 = 0$ it means that the particle does not exist at the walls.

$|\Psi_1|^2$ is maximum at $x = L/2$, it means that the particle exist at the centre of the well.

$$E_1 = \frac{h^2}{8mL^2}$$

This is the energy eigen energy value for ground sate.

For first excited state, n = 2

$$\therefore \quad \psi_2 = A\sin\left(\frac{2\pi}{L}\right)x$$

$\Psi_2 = 0$ for $x = 0$, L/2 and L and Ψ_2 reaches maximum value for $x = L/4$ and 3L/4.

At $x = 0$, L/2 and L, $|\Psi_2|^2 = 0$ it means that the particle does not exist at 0, L/2 and L.

$|\Psi_2|^2$ is maximum at x = L/4 and 3L/4

Energy Eigen values can be calculated by using equation,

$$E_2 = \frac{4h^2}{8mL^2} \quad \text{or} \quad E_2 = 4E_1$$

This is the equation to calculate the energy of the particle in first excited state.

For second excited state, n = 3

$$\psi_3 = A\sin\left(\frac{3\pi}{L}\right)x$$

$\Psi_3 = 0$ for x = 0, L/3, 2L/3 and L, and Ψ_3 reaches maximum value for x = L/6, L/2 and 5L/6.
At x = 0, L/3, 2L/3 and L, $|\Psi_3|^2 = 0$ it means that the particle does not exist at 0, L/3, 2L/3 and L.
$|\Psi_3|^2$ is maximum at x = L/6, L/2 and 5L/6.
Energy Eigen values can be calculated by using equation,

$$E_3 = \frac{9h^2}{8mL^2} \quad \text{or} \quad E_3 = 9E_1$$

This is the equation to calculate the energy of the particle in second excited state.

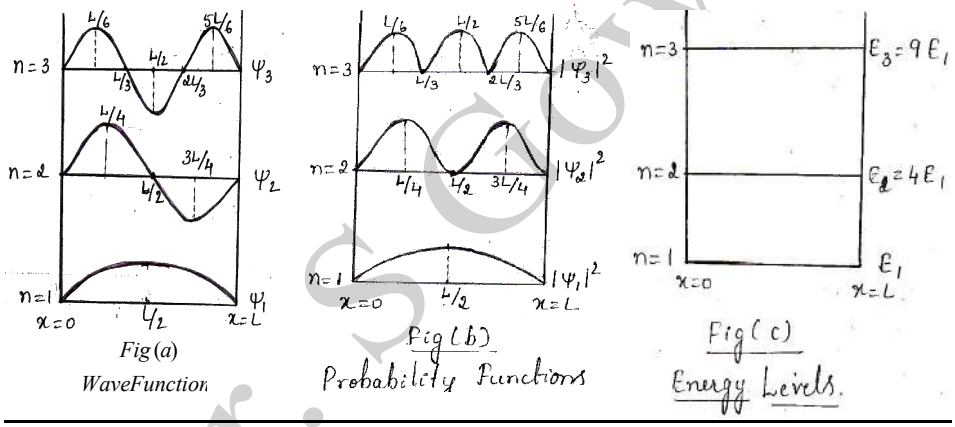


Fig (a) *WaveFunction* Fig (b) Probability Functions Fig (c) Energy Levels.

Problems:

1. Calculate the zero point energy for an electron in a box of width 10Å.

Given, L = 10 Å = 10x10^{-10} m, E_1 or E_0 = ?

w.k.t; $E_n = \frac{n^2h^2}{8mL^2}$

for ground state n = 1

$$E_1 = \frac{h^2}{8mL^2} = \frac{\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(10\times10^{-10}\right)^2} = 6.038\times10^{-20}\,J$$

$$E_1 = \frac{6.038\times10^{-20}}{1.602\times10^{-19}}eV = 0.376\,eV$$

2. An electron is bound in a one dimensional potential well of width 1Å, but of infinite height. Find the energy value for the electron in the ground state.

Given, L = 1 Å = 1x10^{-10} m, E_1 or E_0 = ?

$$w.k.t; \quad E_n = \frac{n^2h^2}{8mL^2}$$

for ground state n = 1

$$E_1 = \frac{h^2}{8mL^2} = \frac{\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(1\times10^{-10}\right)^2} = 6.038\times10^{-18}\,J$$

$$E_1 = \frac{6.038\times10^{-18}}{1.602\times10^{-19}}eV = 37.69\,eV$$

3. An electron is trapped in a one-dimensional box of length 0.1 nm. Calculate the energy required to excite the electron from its ground state to the 2nd excited state.

Given, L = 0.1 nm = 0.1x10^{-9} m , E_1 or E_0 = ?

$$w.k.t; \quad E_n = \frac{n^2h^2}{8mL^2}$$

for ground state n = 1

$$E_1 = \frac{h^2}{8mL^2} = \frac{\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(0.1\times10^{-9}\right)^2} = 6.038\times10^{-18}\,J$$

$$E_1 = \frac{6.038\times10^{-18}}{1.602\times10^{-19}}eV = 37.691\ eV$$

for 2nd excited state n = 3

$$E_3 = \frac{9h^2}{8mL^2} = \frac{9\times\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(0.1\times10^{-9}\right)^2} = 339.219\ eV$$

$$E_3 = 9\times E_1 = 9\times37.691 = 339.219\ eV$$

The energy required to excite the electron from its ground state to the 2nd excited state is,

$$\therefore E = E_3 - E_1 = (339.219 - 37.691)eV$$

$$E = 301.528\,eV$$

4. Calculate the lowest energy of the system consisting of three electrons in a one - dimensional potential box of length 1Å.

Given, L = 1 Å = 1x10^{-10} m , E_1 or E_0 = ?

$$w.k.t; \quad E_n = \frac{n^2h^2}{8mL^2}$$

for lowest energy n = 1

$$E_1 = \frac{h^2}{8mL^2} = \frac{\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(1\times10^{-10}\right)^2} = 1.8114\times10^{-17}\,J$$

$$E_1 = \frac{1.8114\times10^{-17}}{1.602\times10^{-19}}eV = 113.07\ eV$$

5. An electron is constrained to a one-dimensional box of side 1nm. Calculate the first 3-eigen values in electron volt.

Given, $L = 1\,nm = 1x10^{-9}\,m$, E_1 or $E_0 = ?$, $E_2 = ?$, $E_3 = ?$

$$w.k.t;\quad E_n = \frac{n^2h^2}{8mL^2},\quad \text{for ground state n = 1}$$

$$E_1 = \frac{h^2}{8mL^2} = \frac{\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(1\times10^{-9}\right)^2} = 6.038\times10^{-20}\,J$$

$$E_1 = \frac{6.038\times10^{-20}}{1.602\times10^{-19}}eV = 0.376\,eV$$

For 2nd & 3rd excited state n = 2 & n = 3

$$E_2 = 4\times E_1 = 4\times0.376 = 1.504\,eV$$

$$E_3 = 9\times E_1 = 9\times0.376 = 3.384\,eV$$

6. An electron is trapped in one-dimensional infinite potential box of width 0.1nm. Calculate its wavelengths and energies corresponding to first two excited states.

Given, $L = 0.1\,nm = 0.1x10^{-9}\,m$, $E_2 = ?$, $E_3 = ?$, $\lambda_2 = ?$, $\lambda_3 = ?$

$$w.k.t;\quad E_n = \frac{n^2h^2}{8mL^2}$$

for first excited state n = 2

$$E_2 = \frac{4h^2}{8mL^2} = \frac{4\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(0.1\times10^{-9}\right)^2} = 2.415\times10^{-17}\,J$$

$$E_2 = \frac{2.415\times10^{-17}}{1.602\times10^{-19}}eV = 150.76\,eV$$

for second excited state n = 3

$$E_3 = \frac{9h^2}{8mL^2} = \frac{9\left(6.63\times10^{-34}\right)^2}{8\times9.1\times10^{-31}\times\left(0.1\times10^{-9}\right)^2} = 5.434\times10^{-17}\,J$$

$$E_3 = \frac{5.4354\times10^{-17}}{1.602\times10^{-19}}eV = 339.215\,eV$$

w.k.t; $\lambda_n = \frac{2L}{n}$

For first excited state n = 2

$$\lambda_2 = \frac{2L}{2}$$

$$\lambda_2 = L = 0.1\times10^{-9}\,m = 0.1\,nm$$

For second excited state n = 3

$$\lambda_3 = \frac{2L}{3} = \frac{2\times0.1\times10^{-9}}{3}$$

$$\lambda_3 = 6.666\times10^{-11}\,m = 0.066\times10^{-9}\,m = 0.066\,nm$$

7. An electron is trapped in one-dimensional infinite potential box of width 0.15nm. Calculate the amount of energy required to excite an electron from ground state to 3rd excited state.

Given, $L = 0.15\,nm = 0.15x10^{-9}\,m$, $E = E_3 - E_2 = ?$

$w.k.t;\quad E_n = \frac{n^2 h^2}{8mL^2}$

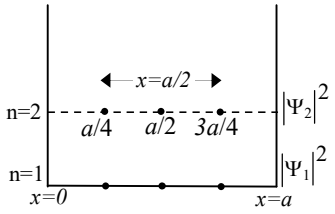
for ground state n = 1 and for third excited state n = 4

$$\therefore E = E_3 - E_2 = \frac{4^2 h^2}{8mL^2} - \frac{h^2}{8mL^2} = \frac{(16-1)h^2}{8mL^2} = \frac{(15)x(6.626\times10^{-34})^2}{8\times9.1\times10^{-31}\times(0.15\times10^{-9})^2} = 9.102\ x10^{-19}$$

$$E = \frac{9.102\ x10^{-19}}{1.602\times10^{-19}} eV = 5.68\ eV$$

8. A quantum particle confined to one dimensional box of width '*a*' in its first excited state. What is the probability of finding the particle over an interval of '*a*/2' marked symmetrically at the centre of the box?

Soln.



The probability of finding the particle can be obtained by

$$p_n = \int_{-\infty}^{\infty} |\psi_n|^2 dx$$

Let a particle be in box of width '*a*',. let p_2 be the probability occupation in the region (*a/2 = 3a/4 - a/4*) symmetrically at the centre therefore the probability of finding the particle in this region in first excited state is obtained by

$$p_2 = \int_{a/4}^{3a/4} |\psi_2|^2 dx = \int_{a/4}^{3a/4} \left|\sqrt{\tfrac{2}{a}}\sin\left(\tfrac{n\pi}{a}\right)x\right|^2 dx = \int_{a/4}^{3a/4} \tfrac{2}{a}\sin^2\left(\tfrac{2\pi}{a}\right)x\,dx$$

$$\therefore p_2 = \tfrac{2}{a}\int_{a/4}^{3a/4} \tfrac{1}{2}\left[1-\cos2\left(\tfrac{2\pi}{a}\right)\right]x\,dx = \tfrac{1}{a}\left[x\Big|_{a/4}^{3a/4} - \tfrac{a}{2x2\pi}\sin2\left(\tfrac{2\pi}{a}\right)x\Big|_{a/4}^{3a/4}\right]$$

$$p_2 = \tfrac{1}{a}\left[3a/4 - a/4 + \tfrac{a}{2x2\pi}\sin2\left(\tfrac{2\pi}{a}\right)3a/4 - \tfrac{a}{2x2\pi}\sin2\left(\tfrac{2\pi}{a}\right)a/4\right] = \tfrac{1}{a}\left[a/2 + 0 - 0\right] = \tfrac{1}{2} = 0.5 = 50\%$$

The probability of occupation in the region *a*/2 at the centre of the box in the 1st excited state is 50%.

9. The ground state energy of an electron in an infinite well is 5.6 MeV. If the width of the well is doubled, calculate the ground state energy.

Given data: $E_1 = 5.6\,meV = 5.6\times10^{-3}\,eV$ *when width a*, $E_1^I = ?$ *when width* $a^I = 2a$,

$$E_1 = \frac{h^2}{8ma^2} \;\&\; E_1^I = \frac{h^2}{8m(2a)^2}$$

$$\frac{E_1^I}{E_1} = \frac{(1/2)^2}{1} = \frac{1}{4}$$

$$\therefore E_1^I = \frac{E_1}{4} = \frac{5.6\times10^{-3}}{4} = 1.4\times10^{-3}\,eV = 1.4\,meV$$

10. An electron is trapped in a 1-D potential well of infinite height and of width of 0.1nm. Calculate the energy required to excite it from its ground state to fifth excited state.

Given data: a = 0.1nm = $0.1\text{x}10^{-9}$m, $E = E_6 - E_1 = ?$

$E = \frac{n^2 h^2}{8ma^2}$ (Note: when n=1 ground state, n-6 for 5th first excited state)

$$Ground\ state \quad E_1 = \frac{1^2 \times \left(6.63 \times 10^{-34}\right)^2}{8 \times 9.1 \times 10^{-31} \times \left(0.1 \times 10^{-9}\right)^2} = 6.038 \times 10^{-18} J = 37.69 eV$$

$For\ 5th\ excited\ state, n = 6$

$$\therefore E_6 = \frac{6^2 \times \left(6.63 \times 10^{-34}\right)^2}{8 \times 9.1 \times 10^{-31} \times \left(0.1 \times 10^{-9}\right)^2} = 2.174 \times 10^{-16} J = 1357 eV$$

$$\therefore E = E_6 - E_1 = 1357 eV - 37.69 eV = 1319.31\, eV$$

Questions:

1. Define wave function? Mention its basic properties.
2. What is a wave function? Explain its physical significance.
3. Discuss the physical interpretation of wave function?
4. Discuss about probability and normalisation condition.
5. Set up time independent Schrödinger's wave equation.
6. Derive Time independent Schrodinger wave equation for a particle moving in one dimension.
7. Explain the significance of wave function and set up time independent Schrödinger's wave equation.
8. Describe Eigen functions and Eigen values.
9. Derive the expressions for eigen values and eigen functions of a particle in one dimensional potential box.
10. Solve Schrödinger wave equation for allowed energy values in case of a particle in a potential box.
11. Obtain the expression for normalised wave function for a particle in one dimensional potential box.
12. Assuming the time independent Schrödinger's wave equation, discuss the solution for a particle in one dimensional potential well of infinite height,
13. Discuss and mention the energy eigen values, eigen wave functions and probability densities for a particle in 1-dimensional box for atleast 3 states.
14. Discuss the energies of a free particle using time independent Schrodinger equation.

Problems:

1. An electron is confined to a one dimensional box of width 1 nm. Calculate the first three Eigen values in eV.
2. An electron is bound in a one dimensional box of width 4 X 10^{-10} m. compute the energy and de-Broglie wavelength of ground and first excited states.
3. Find the energy of an electron in the ground state, when it is trapped in an infinite potential well of width 2Å.
4. An electron is bound in one dimensional potential well of width 1Å but of infinite height. Find its energy values in ground state and first two exited states.
5. Estimate the time spent by an atom in the excited state during the excitation and de-excitation processes, when a spectral line of wavelength 546 nm and width 10^{-14} m is emitted.
6. An electron is confined to a box of length 10^{-9} m, calculate the minimum uncertainty in its velocity.
7. The position and momentum of 1 keV electron are simultaneously determined. If it position is located within 1Å, find the uncertainty in the determination of its momentum.
8. A spectral line of wavelength 4000 Å has a width of 8×10^{-5}Å. Evaluate the minimum time spent by the electrons in the upper energy state between the excitation and de-excitation processes.
9. The inherent uncertainty in the measurement of time spent by Iridium 191 nuclei in the excited state is found to be 1.4×10^{-10}s. Estimate the uncertainty that results in its energy in eV in the excited state.
10. An electron is bound in one dimensional potential well of width 0.18 nm. Find the energy value in eV of the second excited state.
11. The first excited state energy of an electron in an in finite well is 240 eV. What will be its ground state energy when the width of the potential well is doubled?
12. A quantum particle confined to one–dimensional box of width *'a'* is in its first excited state. What is the probability of finding the particle over an interval of *a*/2 marked symmetrically at the center of the box.

Free Particle

Energy Eigen values for a free particle:

Free particle means, it is not under the influence of any kind of field or force. Thus it has zero potential, i.e., V = 0.

Hence Schrödinger's equation becomes,

$$\frac{d^2\psi}{dx^2} = -\frac{8\pi^2 m}{h^2}(E-V)\psi = 0$$

Since V = 0, $$\frac{d^2\psi}{dx^2} = -\frac{8\pi^2 m}{h^2}E\psi = 0$$

The above equation holds good for a particle for which the potential v=0 over the entire space (No boundaries at all).

We know that in the case of particle in an infinite potential well, the condition V=0 holds good only over a infinite width 'L' and outside region, V= ∞ ,

Since for the free particle, V = 0 holds good everywhere, we can extend the case of particle in an infinite potential well to the free particle's case, by treating the width of the well to be infinity, i.e., by allowing L = ∞,

We have the equation for energy Eigen values for a particle in an infinite potential well as,

$$E = \frac{n^2 h^2}{8mL^2}$$

Where, n = 1, 2, 3, ...

Rearranging the above equation, we have,

$$n = \frac{2L}{h}\sqrt{2Em}$$

Here, we see that for a particle with constant energy E but confined in the well, n depends mainly on 'L'. Hence as L→ ∞, n→∞. If the particle is no more confined in any sort of well but free, at that time it also follows that n = ∞, which essentially means that a free particle can have any energy i.e., the energy Eigen values or the possible values of energy are infinite in number. Keeping the energy level representation in the mind, we say that the permitted energy values are continues. All these mean, there is no discreteness in the allowed energy values. In other word, there is no quantization of energy in case of a free particle and the problem is dealt in classical mechanics. Thus a free particle is a classical entity.

https://www.youtube.com/watch?v=tlM9vq-bepA

https://www.youtube.com/watch?v=v9DPzMoWpc0

https://www.youtube.com/watch?v=8l4x4vbMP0c

MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2022-23

Module – 3 : Quantum Computing

Notes

Syllabus:

Quantum Computing: **8 hrs**

Principles of Quantum Information & Quantum Computing:

Introduction to Quantum Computing, Moore's law & its end, Differences between Classical & Quantum computing. Concept of qubit and its properties. Representation of qubit by Bloch sphere. Single and Two qubits. Extension to N qubits.

Dirac representation and matrix operations:

Matrix representation of 0 and 1 States, Identity Operator I, Applying I to|0⟩and |1⟩ states, Pauli Matrices and its operations on |0⟩and |1⟩states, Explanation of i) Conjugate of a matrix and ii) Transpose of a matrix. Unitary matrix U, Examples: Row and Column Matrices and their multiplication (Inner Product), Probability, and Quantum Superposition, normalization rule. Orthogonality, Orthonormality. Numerical Problems.

Quantum Gates:

Single Qubit Gates: Quantum Not Gate, Pauli – X, Y and Z Gates, Hadamard Gate, Phase Gate (or S Gate), T Gate Multiple Qubit Gates: Controlled gate, CNOT Gate, (Discussion for 4 different input states). Representation of Swap gate, Controlled -Z gate, Toffoli gate.

Pre requisites: Matrices

Self-learning: Moore's law

Module - 3

Quantum Computing

Introduction

Quantum computing is a rapidly-emerging field focused on the development of computer technologies centered on the principles of Quantum Physics. Quantum Physics explains the nature and behaviour of energy and matter on the quantum (atomic and subatomic) scale. Elementary particles such as protons, neutrons and electrons can exist in two or more states at a time. This fundamental behaviour is utilized in designing the quantum computation processing units and in fact it is more efficient than classical computation

Quantum computing uses a combination of bits to perform specific computational tasks with greater efficiency than their classical counterparts. Even though quantum computers are not going to replace classical computers, quantum technology is significantly changing the way the world operates. The quantum computer gains much of its processing power through the ability for bits to be in multiple states simultaneously. They can perform tasks using a combination of 1's, 0's and both 1 and 0 at a time

Brief History

In 1981, Paul Benioff at Argonne National Labs came up with the idea of a computer that operates with quantum mechanical principles. In 1984, David Deutsch of Oxford University provided the critical idea behind quantum computing research and the possibility of designing a computer that is based exclusively on quantum rules.

The Essential Elements of Quantum Theory

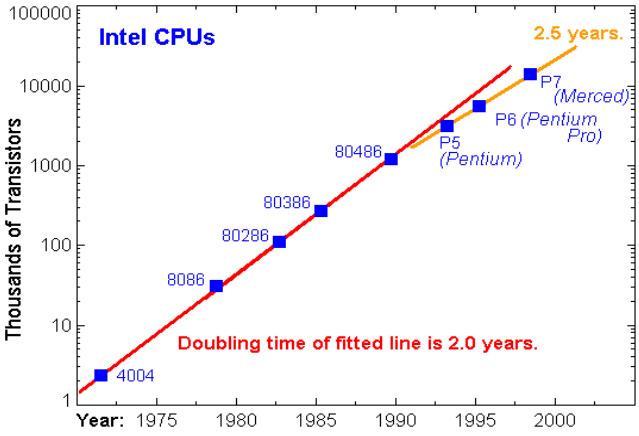
- Energy values are discrete units.
- Elementary particles may behave like particles or waves.
- The movement of elementary particles is inherently random and, thus, unpredictable.
- The simultaneous measurement of two complementary values - such as the position and momentum of a particle - is imperfect. The more precisely one value is measured, the more flawed the measurement of the other value will be.

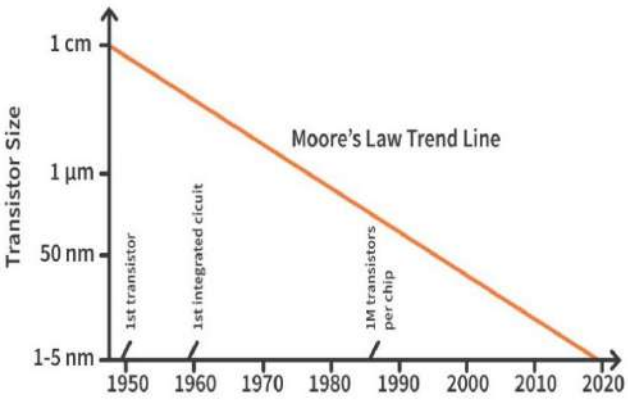
Moore's law & its end

In 1965, Gordon E. Moore—co-founder of Intel—postulated that *"the number of transistors that can be packed into a given unit of space will double about every eighteen months"*. This is also known as Moore's Law

Gordon Moore did not call his observation as "Moore's Law," nor did he set out to create a "law". Moore made this statement based on noticing emerging trends in chip manufacturing at the semiconductor industry. Eventually, Moore's insight became a prediction, which in turn became the golden rule known as Moore's Law.

Moore's Law implies that computers, machines that run on computers, and computing power all become smaller, faster, and cheaper with time, as transistors on integrated circuits become more efficient.
Here is a graphic representation for microprocessors





Impact of Moore's Law on Computing

Moore's Law has had a direct impact on the progress of computing power. What this means specifically, is that transistors in integrated circuits have become faster. Transistors conduct electricity, which contain carbon and silicon molecules that can make the electricity run faster across the circuit. The faster the integrated circuit conducts electricity, the faster the computer operates.

Is Moore's Law Coming to an End?

The electronic industry for computers grows hand-in-hand with the decrease in size of the integrated circuits. This miniaturization is necessary to increase computational power, that is, the number of floating-point operations per second (FLOPS) a computer can perform. In 1950's, electronic computers were capable of performing approximately 10^3 FLOPS while present supercomputers have a power greater than 10^{13} FLOPS. According to Moore's law, the number of transistors that may be placed on a single integrated circuit chip doubles approximately every 18 – 24 months. The present limit is approximately **10^8 transistors** per chip and the typical size of circuit components is of the order of 100 nano meters. That means, we have reached the atomic size for storing a single bit of information and quantum effects have become unavoidably dominant.

Taking all these factors into consideration, it is necessary to look for alternative ways of computing outside of the electrons and silicon transistors. One such alternative is quantum computing.

Quantum computers are based on quantum bits (qubits) and use quantum effects like *superposition* and *entanglement* to their benefit, hence overcoming the miniaturization problems of classical computing.

Comparison of Classical and Quantum Computing

Classical computing relies on principles of Boolean algebra. Data must be processed in an exclusive binary state at any point in time; either 0 (off / false) or 1 (on / true). The millions of transistors and capacitors at the heart of computers can only be in one state at any point. In addition, there is still a limit as to how quickly these devices can be made to switch states. As we progress to smaller and faster circuits, we begin to reach the physical limits of materials and the limitations for classical laws of physics to apply

The quantum computer operates with a two-mode logic gate. In a quantum computer, a number of elemental particles such as electrons or photons can be used. Each particle is given a charge or polarization, acting as a representation of 0 and/or 1. Each particle is called a *quantum bit*, or *qubit*. The nature and behaviour of these particles form the basis of quantum computing. The two most relevant aspects of quantum physics are the principles of superposition and entanglement.

Differences between classical and quantum computing

Comparison key	Classical computer	Quantum computer
Basis of computing	Large scale multipurpose computer based on classical physics	High speed computer based on quantum mechanics
Information storage	Bit-based information storage using voltage/charge	Quantum bit-based information storage using electron spin or polarization
Bit values	Bits having a value of either 0 or 1 can have a single value at any instant	Qubits have a value of 0, 1 or sometimes linear combination of both, (a property known as superposition)
Number of possible states	The number of possible states is 2 which is either 0 or 1	The number of possible states is infinite since it can hold combinations of 0 or 1 along with some complex information
Output	Deterministic (repetition of computation on the same input gives the same output)	Probabilistic (repetition of computation on superposed states gives probabilistic answer)
Gates used for processing	Logic gates (AND, OR, NOT, etc.)	Quantum gates (X, Y, Z, H, CNOT etc.)
Operations	Operations use Boolean Algebra	Operations use linear algebra and are represented with unitary matrices
Circuit implementation	Circuit implemented in macroscopic technologies (e.g. CMOS) that are fast and scalable	Circuits implemented in microscopic technologies (e.g. nuclear magnetic resonance) that are slow and delicate

Concept of qubit and its properties

From bits to qubit

Bit: A digital computer stores and processes information using bits, which can be either 0 or 1. Physically, a bit can be anything that has two distinct configurations: one represented by "0", and the other represented by "1". It could be a light bulb that is on or off, a coin that is heads or tails, or any other system with two distinct and distinguishable possibilities. In modern computing and communications, bits are represented by the absence or presence of an electrical signal, encoding "0" and "1" respectively

Qubit is the physical carrier of quantum information. It is the quantum version of a bit, and its quantum state can be written in terms of two levels, labelled $|0\rangle$ and $|1\rangle$. $|\ \rangle$ this notation is known as 'ket' notation and $\langle\ |$ is known as 'brac' notation. Both are together called as **Dirac notations**
'Ket' is analogous to a column vector.
They are also called basis vectors and represented by two-dimensional column vectors as follows

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

The qubit can be in any one of the two states as well as in the superposed state simultaneously

In quantum computation two distinguishable states of a system are needed to represent a bit of data. For example, two states of an electron orbiting a single atom is shown in the following figure. Spin up is taken as $|1\rangle$ and spin down is taken as $|0\rangle$. Similarly ground state energy level is $|0\rangle$ and excited state level is $|1\rangle$

Qubit represented by two electronic levels in an atom

This is the abstract notion of a qubit. The quantum computers actually use a physical type of qubit called a *superconducting qubit* is made from superconducting materials (of course, there are other approaches also to build the qubits)

NOTE:
In quantum computing the vectors are members of complex vector space[#]. Each member of this space is represented as column vector of n dimensions with single 1 at the location corresponding to a particular basis vector.
It is as follows

$$|0\rangle = \begin{bmatrix}1\\0\\0\\\cdot\\\cdot\\0\end{bmatrix} \quad |1\rangle = \begin{bmatrix}0\\1\\0\\\cdot\\\cdot\\0\end{bmatrix} \quad |2\rangle = \begin{bmatrix}0\\0\\1\\\cdot\\\cdot\\0\end{bmatrix} \ldots\ldots \quad |N-1\rangle = \begin{bmatrix}0\\0\\0\\\cdot\\\cdot\\1\end{bmatrix}$$

[#] Complex vector space is a vector space which contains complex numbers

Here we use only two dimensions (or only two sets). Hence we write as

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

Superposition of two states

The difference between qubits and classical bits is that a qubit can be in a linear combination (superposition) of the two states $|0\rangle$ and $|1\rangle$. For ex, if $\propto$ and β are the probability amplitudes of electron in ground state (ie, in $|0\rangle$ state) and in excited state (ie, in $|1\rangle$ state) then the linear combination of two states is

$$|\psi\rangle = \alpha\,|0\rangle + \beta|1\rangle$$

The numbers α and β are complex but due to normalization conditions

$$|\propto|^2 + |\beta|^2 = 1.$$

Here $|\propto|^2$is the probability of finding $|\psi\rangle$ in state $|0\rangle$ and $|\beta|^2$is the probability of finding $|\psi\rangle$ in state $|1\rangle$. So, we have to keep in mind that when a qubit is measured, it only gives either '0' or '1' as the measurement result – probabilistically

Consider the following example of qubit representation

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$
$$\therefore \propto = \frac{1}{\sqrt{2}} \; and \;\; \beta = \frac{1}{\sqrt{2}}$$
$$|\propto|^2 = |\beta|^2 = \frac{1}{2}$$

This means that with 50% probability the qubit will be found in $|0\rangle$ state as well as in $|1\rangle$ state. The superposed states are also called as **space states** where as $|0\rangle$ and $|1\rangle$ are called **basis states**.

Properties of qubits

1. Qubits make use of discrete energy state particles such as electrons and photons
2. Qubits exists in two quantum state $|0\rangle$ and $|1\rangle$ or in a linear combination of both states. This is known as superposition. This property allows for *exponentially many logical states* at once (and no classical computer can achieve it)
3. Unlike classical bits, qubit can work with the overlap of both 0 & 1 states. For ex, a 4-bit register[1] can store one number from 0 to 15 (because of $2^n = 2^4$=16), but 4-qubit register can store all 16 numbers
4. When the qubit is measured, it collapses to one of the two basis states $|0\rangle$ or $|1\rangle$
5. Quantum entanglement and quantum tunnelling are two exclusive properties of qubit
6. State of the qubits is represented using Bloch sphere

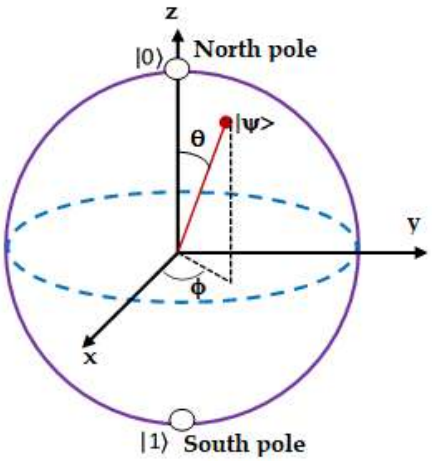
[1] Register – is a group of flip-flops. Its basic function is to hold information within a digital system so as to make it available to the logic units during the computing process. However, a register may also have additional capabilities associated with it.

After studying the physics of qubits it is now time to look at the mathematics of qubits. Let us start with the representation of qubit using Bloch sphere in a vector space. Later on we proceed towards single qubit, multi qubit, tensor operation, operators and matrix representation

NOTE: Vector space is a set of elements (or vectors) which are added together or multiplied by real numbers. Addition of two vectors or multiplication of a vector by a scalar is satisfied here. It should not be confused with vector field

Bloch sphere

Bloch sphere is an imaginary sphere which is used to represent pure single-qubit states as a point on its surface. It has unit radius. Its North Pole and South Pole are selected to represent the basis states namely $|0\rangle$ and $|1\rangle$. North Pole represents $|0\rangle$ (say spin up↑) and South Pole represents $|1\rangle$ (say spin down ↓). All other points on the sphere represent superposed states (ie, state space). Bloch sphere allows the state of a qubit to be represented in spherical coordinates (ie, r, θ and ϕ). It is as follows



The state qubit $|\psi\rangle$ on the Bloch sphere makes an angle θ with z-axis and its projection (azimuth) makes angle ϕ with x-axis as shown. It is clear from the fig that $0 < \theta < \pi$ and $0 < \phi < 2\pi$. $|\psi\rangle$ is represented as

$$|\psi\rangle = \alpha\,|0\rangle + \beta|1\rangle$$

It can be proved that

$$|\Psi\rangle = cos\frac{\theta}{2}|0\rangle + e^{i\phi}sin\frac{\theta}{2}|1\rangle - - - (1)$$

Using this equation we can represent $|\psi\rangle$ for different θ and ϕ as follows

Case-1: let $\theta = 0$ and $\phi = 0$, then eq (1) becomes

$$|\Psi\rangle = cos\ 0|0\rangle + e^{i0}sin0|1\rangle = |0\rangle + 0$$
$$\therefore |\Psi\rangle = |0\rangle$$

Case-2: let $\theta = \pi$ and $\phi = 0$, then eq (1) becomes

$$|\Psi\rangle = cos\ \frac{\pi}{2}|0\rangle + e^{i0}sin\frac{\pi}{2}|1\rangle = 0 + |1\rangle$$

$$\therefore |\Psi\rangle = |1\rangle$$

Case-3: let $\theta = \pi/2$ and $\phi = 0$, then eq (1) becomes

$$|\Psi\rangle = cos\ \frac{\pi}{4}|0\rangle + e^{i0} sin\frac{\pi}{4}|1\rangle$$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

$$|\Psi\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

Case-4: let $\theta = \pi/2$ and $\phi = \pi$, then eq (1) becomes

$$|\Psi\rangle = cos\ \frac{\pi}{4}|0\rangle + e^{i\pi} sin\frac{\pi}{4}|1\rangle$$

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

$$|\Psi\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

In the above discussion we have represented only single qubit state. Bloch sphere is a nice visualization of single qubit states.

Multiple Qubits

Single qubits are interesting, but individually they offer no computational advantage. We will now look at how to represent multiple qubits, and how these qubits can interact with each other.

Two qubits

Consider two qubits. They can be in any one of four possible states represented as $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. The interaction among these qubits is described by creating a new vector space[2] using a special kind of operation called *tensor product* or *Kronecker product*. It is as follows

Let U and V are two 2-d vectors given as

$$U = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, V = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

Their tensor product is

$$U \otimes V = \begin{bmatrix} x_1 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \\ y_1 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \end{bmatrix}$$

[2] Vector space is a set of elements (or vectors) which are added together or multiplied by real numbers or scalars. Addition of two vectors or multiplication of a vector by a scalar is satisfied here. It should not be confused with vector field

$$U \otimes V = \begin{bmatrix} x_1x_2 \\ x_1y_2 \\ y_1x_2 \\ y_1y_2 \end{bmatrix}$$

Based on this we can write |00⟩ as follows

$$|0\rangle \otimes |0\rangle = \begin{bmatrix} 1\begin{bmatrix}1\\0\end{bmatrix} \\ 0\begin{bmatrix}1\\0\end{bmatrix} \end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

Sometimes we avoid the symbol ⊗ and write directly as

$$|00\rangle = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

Similarly

$$|01\rangle = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad |10\rangle = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \text{ and } |11\rangle = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

The state qubit is (ie, linear combination of these four)

$$|\Psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

For 2 qubit system we have 4 complex amplitudes namely α_{00}, α_{01}, α_{10} and α_{11}. According to normalization condition

$$|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$$

Similarly if there are 3 qubits there will be 8 complex amplitudes and in general for *n* qubits we will have 2^n complex amplitudes. This means that a basis state is represented by a number 0 to 2^{n-1}. The superposition state is represented as

$$|\Psi\rangle = \sum_{x=0}^{2^{n-1}} \propto_x |x\rangle$$

Qubit has two quantum states similar to the classical binary states. The qubit can be in any one of the two states as well as in the superposed state simultaneously.

Dirac representation and Matrix operations

Matrix representation of 0 and 1 states

In Quantum mechanics, Brac-Ket notation is a standard notation for describing quantum states. The notation | ⟩ is known as 'ket' notation and ⟨ | is known as 'brac' notation. Both are together called as **Dirac notations.**

The ‘ket’ vector typically represented as a column vector and ‘brac’ vector typically represented as a row vector as follows

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix} - - - - - - \text{ket notations}$$

$$\langle 0| = [1 \quad 0] \qquad \langle 1| = [0 \quad 1] - - - - - \text{ brac notations}$$

Hence, any arbitrary state can be represented as

$$|\psi> = \begin{bmatrix}\alpha\\\beta\end{bmatrix} \quad or \quad |\psi\rangle \; = \; \alpha\,|0\rangle \; + \beta\,|1\rangle$$

Some of the properties of these notations are

i. Addition of two kets gives another ket (commutative)

$$|\mathrm{A}\rangle + |B\rangle = |C\rangle = |B\rangle + |A\rangle$$

ii. Addition of kets obeys associative property

$$|\mathrm{A}\rangle + (|B\rangle + |C\rangle) = (|A\rangle + |B\rangle) + |C\rangle$$

iii. If c_1 and c_2 are scalars or a complex numbers and |A ⟩ is a ket then

$$(c_1 + c_2)|A\rangle = c_1|A\rangle + c_2|A\rangle$$

iv. In a complex vector space for every ket there is unique brac. Brac is the Hermitian[#] conjugate of the ket.

$$\text{If } |A\rangle = \begin{bmatrix}A_1\\A_2\end{bmatrix} \text{ then } \langle A| = [A_1^* \quad A_2^*]$$

v. Bracs are useful in calculating probability amplitudes.
For ex, the probability amplitude of |1⟩ is β which can be calculated as follows

$$\langle 1||\Psi\rangle = \langle 1|\alpha\,|0\rangle \; + \langle 1|\beta\,|1\rangle$$
$$\langle 1||\Psi\rangle = \alpha\langle 1|\,|0\rangle \; + \beta\,\langle 1||1\rangle$$
$$\langle 1||\Psi\rangle = \alpha[0 \quad 1]\begin{bmatrix}1\\0\end{bmatrix} + \beta[0 \quad 1]\begin{bmatrix}0\\1\end{bmatrix}$$
$$\langle 1|\Psi\rangle = \alpha \times 0 + \beta \times 1$$
$$\langle 1|\Psi\rangle = \beta$$

Similarly

$$\langle 0|\Psi\rangle = \alpha$$

vi. If ⟨U| and ⟨V| are two bracs then

[#] Will be discussed later

$$\langle U| + \langle V| = \langle U + V|$$

Operators and matrices

An operator is a mathematical rule that transform a given function into another function.
Ex:

i. $\sqrt{4} = 2$. Here $\sqrt{}$ is a square root operator. It transforms 4 to 2
ii. $D = \frac{d}{dx}$ is a differentiate operator. It transforms $2x^3$ to $6x^2$

Similar to this we have the following example. In this case operator 'A' transforms the vector $|a\rangle$ to another vector $|b\rangle$

$$\hat{A}\ |a\rangle\ =\ |b\rangle$$

There are different types of operators like Linear operator, Identity operator, Null operator, Inverse operator, Singular & non-singular operator etc.

Identity operator 'I'

The identity operator is an operator which, operating on a function, leaves the function unchanged i.e.

$$I\ |a\rangle\ =\ |a\rangle$$

It is given in matrix form by

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This is also called as identity matrix. There will be no change when I operates on either $|0\rangle$ state or $|1\rangle$ state. It is explained as follows

$$I\ |0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\therefore I\ |0\rangle\ =\ |0\rangle$$

Similarly

$$I\ |1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\therefore I\ |1\rangle\ =\ |1\rangle$$

Identity matrix acts as number 1. It is always a square matrix.

Conjugate matrices

If the elements in a matrix A are complex numbers, then the matrix obtained by the corresponding conjugate complex elements is called the **conjugate** of A and is denoted by A^*. For ex

$$\text{If } A = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \text{ then } \quad A^* = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$\text{If } A = \begin{bmatrix} i & 2i+1 \\ -i & 1 \end{bmatrix} \text{ then } A^* = \begin{bmatrix} -i & -2i+1 \\ i & 1 \end{bmatrix}$$

$$\text{If } A = \begin{bmatrix} 1 & 2i \\ 4i+1 & 0 \end{bmatrix} \text{ then } A^* = \begin{bmatrix} 1 & -2i \\ -4i+1 & 0 \end{bmatrix}$$

$$\text{If } A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \text{ then } A^* = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$$

Transpose matrices

If columns and rows of a matrix A are interchanged then the resultant matrix is **transpose** of A and represented as A^T. For ex,

$$\text{If } A = \begin{bmatrix} 0 & 1 \\ -i & 0 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 0 & -i \\ 1 & 0 \end{bmatrix}$$

$$\text{If } A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

$$\text{If } A = \begin{bmatrix} 1 & 2i \\ 4i+1 & 0 \end{bmatrix} \text{ then } A^T = \begin{bmatrix} 1 & 4i+1 \\ 2i & 0 \end{bmatrix}$$

Hermitian matrices

The transpose of complex conjugate of a matrix is known as Hermitian operator (also called as adjoint operator) and the resultant matrix is known as **Hermitian matrix.** It is represented by $A^\dagger$

Let A be a matrix, A^* be its complex conjugate and A^{*T} is its transpose then its **Hermitian** matrix is

$$A^\dagger = A^{*T}$$

Ex:

$$\text{If } A = \begin{bmatrix} 1 & 2i \\ 4i+1 & 0 \end{bmatrix} \text{ then } A^* = \begin{bmatrix} 1 & -2i \\ -4i+1 & 0 \end{bmatrix}$$

$$A^\dagger = \begin{bmatrix} 1 & -4i+1 \\ -2i & 0 \end{bmatrix}$$

Unitary matrices

Matrix A is said to be unitary if it produces an identity matrix *I* when multiplied by its conjugate transpose

$$AA^\dagger = I$$

In other words, A is a unitary matrix if its conjugate transpose is equal to its reciprocal, ie

$$A^\dagger = \frac{I}{A} = \frac{1}{A} = A^{-1}$$

we can show that $A = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$ is a unitary matrix

Inner product

Introduction

Let $U = x_1 i + y_1 j + z_1 k$ and $V = x_2 i + y_2 j + z_2 k$ be the two vectors in real space then their dot product is

$$U.V = x_1 x_2 + y_1 y_2 + z_1 z_2$$

If U = V then

$$U.U = |U|^2 = x_1^2 + x_2^2 + x_3^2$$

The length of the resultant vector is $|U| = \sqrt{U.U} = \sqrt{x_1^2 + y_1^2 + z_1^2}$

In matrix form U and V are written as

$$U = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \text{ and V} = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

And the dot product is written as

$$U.V = [x_1 y_1 z_1] \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = U^T V$$

This dot product is also called as *inner product.* In real space inner product is same as dot product of two vectors and it finally gives a scalar quantity.

In quantum computing the vectors are the members of *complex space* and the *inner product* gives a complex number

Definition of inner product

The inner product of two vectors U and V in the complex space is a function that takes U and V as inputs and produces a complex number as output

In terms of Dirac notation, the inner product is given as

$$\langle U|V\rangle = c$$

Let $|U\rangle = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $|V\rangle = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ be the two vectors. Their inner product is written as $\langle U|V\rangle$

But $\langle U|$ is equal to conjugate transpose of $|U\rangle$

$$ie, \langle U| = |U^*\rangle^{-1} = |U\rangle^{\dagger} = [x_1^* \quad y_1^*]$$

$$\therefore \langle U|V\rangle = [x_1^* \quad y_1^*] \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = x_1^* x_2 + y_1^* y_2$$

The square root of the inner product of a vector with itself is also called as ***norm*** **o**r the *length* of the vector. It is given by

$$|U| = \sqrt{\langle U|U\rangle}$$

Ex: Find the inner product of $|\boldsymbol{U}\rangle = \begin{bmatrix} \boldsymbol{3+i} \\ \boldsymbol{4-i} \end{bmatrix}$ **and** $|\boldsymbol{V}\rangle = \begin{bmatrix} \boldsymbol{3i} \\ \boldsymbol{4} \end{bmatrix}$

First we shall find the conjugate transpose of $|U\rangle$

$$|U^*\rangle = \begin{bmatrix} 3-i \\ 4+i \end{bmatrix}$$

$$|U\rangle^\dagger = [3-i \quad 4+i]$$

$$\therefore \langle U| = |U\rangle^\dagger = [3-i \quad 4+i]$$

$$\langle U|V\rangle = [3-i \quad 4+i]\begin{bmatrix} 3i \\ 4 \end{bmatrix}$$

$$\langle U|V\rangle = (3-i)\times 3i + (4+i)\times 4$$

$$\langle U|V\rangle = 9i + 3 + 16 + 4i$$

$$\langle U|V\rangle = 13i + 19$$

Ex: Find the inner product of $|\boldsymbol{A}\rangle = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{ib} \end{bmatrix}$ ***with itself***

First we shall find the conjugate transpose of |A⟩

$$|A^*\rangle = \begin{bmatrix} a \\ -ib \end{bmatrix}$$

$$|U\rangle^\dagger = [a \quad -ib]$$

$$\therefore \langle A| = [a \quad -ib]$$

$$\langle A|A\rangle = [a \quad -ib]\begin{bmatrix} a \\ ib \end{bmatrix}$$

$$\langle A|A\rangle = a^2 + (-ib)(ib)$$

$$\langle A|A\rangle = a^2 + b^2$$

Ex: find the norm of $|\boldsymbol{U}\rangle = \begin{bmatrix} \boldsymbol{1-i} \\ \boldsymbol{2} \end{bmatrix}$

$$|U| = \sqrt{\langle U|U\rangle}$$

$$|U| = \sqrt{[1+i \quad 2]\begin{bmatrix} 1-i \\ 2 \end{bmatrix}}$$

$$|U| = \sqrt{(1+i)(1-i) + 2\times 2} = \sqrt{1+1+4} = \sqrt{6}$$

Orthogonality

If the inner product of two vectors is equal to 0 then they are said to be ***orthogonal*** *(or perpendicular)* to each other

If $\langle U|V\rangle = 0$ then $|U\rangle$ and $|V\rangle$ are perpendicular.

Consider,

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

Then

$$\langle 0|1\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = 0$$

Hence $|0\rangle$ is perpendicular to $|1\rangle$

The most important property of the inner product of a vector with itself is equal to one

$$\text{ie, } \langle\psi|\psi\rangle = 1$$

This is known as normalization condition. The physical significance of normalization is that the "probability amplitude" of the quantum system is1

Orthonormality

If each element of a set of vectors is normalized and the elements are orthogonal with respect to each other, we say the set is ***orthonormal*** (ortho + normalization = orthonormalization)

Consider the set

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix} \qquad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

$\langle 0|0\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = 1 + 0 = 1$ normalized

$\langle 0|1\rangle = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = 0 + 0 = 0$ orthogonal

$\langle 1|1\rangle = \begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = 0 + 1 = 1$ normalized

$\langle 1|0\rangle = \begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = 0 + 0 = 0$ orthogonal

Hence set of $|0\rangle$ and $|1\rangle$ is orthonormal

Pauli Matrices

These are the 2 × 2 complex matrices introduced by Pauli in order to account for the interaction of the spin with an external electromagnetic field. They are given by

$$\sigma_1 = \sigma_j = X = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$$

$$\sigma_2 = \sigma_k = Y = \begin{bmatrix}0 & -i\\i & 0\end{bmatrix}$$

$$\sigma_3 = \sigma_l = Z = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$$

NOTE: X, Y and Z are also called as X – gate, Y- gate and Z- gate

Properties of Pauli matrices

The most important property of Pauli matrices is that square of all the three matrices gives an identity matrix I. For ex,

$$\sigma_1^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\therefore \sigma_1^2 = I$$

In general

$$\sigma\sigma^\dagger = 1$$

$$\sigma^\dagger = \frac{1}{\sigma} = \sigma^{-1}$$

So, they are unitary

Another property of Pauli matrices is that they are Hermitian. Let A be a matrix, A^* be its complex conjugate and $A^\dagger$ [3] is its transpose. If A = $A^\dagger$then the matrix is Hermitian. For ex,

$$\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$${\sigma_2}^* = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$$

$${\sigma_2}^\dagger = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$\therefore {\sigma_2}^\dagger = \sigma_2$$

Operation of Pauli Matrices on 0 and 1 states

Three Pauli matrices X, Y and Z act on basis states $|0\rangle$ and $|1\rangle$ as follows

i. X operating on $|0\rangle$ and $|1\rangle$

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$$

$$X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$$

Since X inverts each input (ie, $|0\rangle$ becomes $|1\rangle$ and $|1\rangle$ becomes$|0\rangle$) it is also called as *bit-flip* gate
If a superposed qubit goes through X gate, the result will be

$$X|\Psi\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \alpha|1\rangle + \beta|0\rangle$$

So,

$$X|\Psi\rangle = \alpha|1\rangle + \beta|0\rangle$$

[3] Transpose means convert rows into column and columns into row

ii. Y operating on $|0\rangle$ and $|1\rangle$

$$Y|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0\times 1 + (-i)\times 0 \\ i\times 1 + 0\times 0 \end{bmatrix} = \begin{bmatrix} 0+0 \\ i+0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i\begin{bmatrix} 0 \\ 1 \end{bmatrix} = i|1\rangle$$

$$Y|1\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0\times 0 + (-i)\times 1 \\ i\times 0 + 0\times 1 \end{bmatrix} = \begin{bmatrix} 0-i \\ 0+0 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix} = -i\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -i|0\rangle$$

So,

$$Y|0\rangle = i|1\rangle$$

Similarly

$$Y|1\rangle = -i|0\rangle$$

If a superposed qubit goes through Y gate, the result will be

$$Y|\Psi\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0\times\alpha + (-i)\times\beta \\ i\times\alpha + 0\times\beta \end{bmatrix} = \begin{bmatrix} -i\beta \\ i\alpha \end{bmatrix} = -i\beta|0\rangle + i\alpha|1\rangle$$

So,

$$Y|\Psi\rangle = i\alpha|1\rangle - i\beta|0\rangle$$

iii. Z operating on $|0\rangle$ and $|1\rangle$

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1\times 1 + 0\times 0 \\ 0\times 1 + (-1)\times 0 \end{bmatrix} = \begin{bmatrix} 1+0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$$

$$Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1\times 0 + 0\times 1 \\ 0\times 0 + (-1)\times 1 \end{bmatrix} = \begin{bmatrix} 0+0 \\ 0-0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 11 \end{bmatrix} = -|1\rangle$$

So,

$$Z|0\rangle = |0\rangle$$

$$Z|1\rangle = -|1\rangle$$

If a superposed qubit goes through Z gate, the result will be

$$Z|\Psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1\times\alpha + 0\times\beta \\ 0\times\alpha + (-1)\times\beta \end{bmatrix} = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix} = \alpha|0\rangle - \beta|1\rangle$$

So,

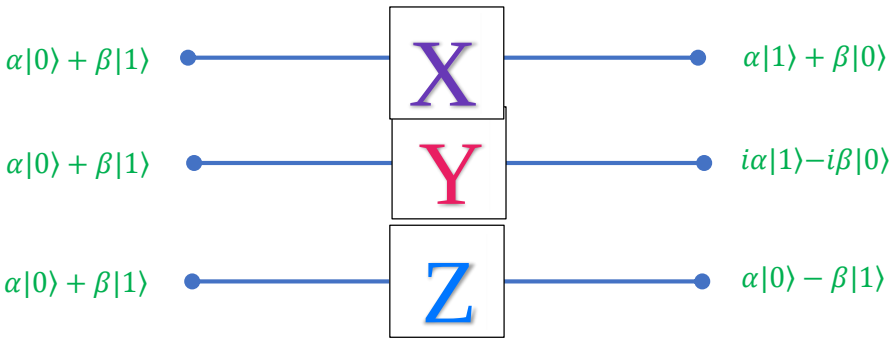
$$Z|\Psi\rangle = \alpha|0\rangle - \beta|1\rangle$$

This is also called *phase-flip* gate

The truth tables for X, Y and Z gates are as follows

X- gate		**Y-gate**		**Z-gate**	
Input	**Output**	**Input**	**Output**	**Input**	**Output**
$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$i\|1\rangle$	$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$-i\|0\rangle$	$\|1\rangle$	$-\|1\rangle$
$\alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|1\rangle + \beta\|0\rangle$	$\alpha\|0\rangle + \beta\|1\rangle$	$i\alpha\|1\rangle - i\beta\|0\rangle$	$\alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|0\rangle - \beta\|1\rangle$

Symbolically these gates are represented as follows



Limitation of Pauli matrices or Pauli gates

Using only the Pauli-gates it is impossible to move our initialized qubit to any state other than $|0\rangle$ or $|1\rangle$, i.e. we cannot achieve superposition. This means we can see no behaviour different to that of a classical bit. To create more interesting states we need more gates

Quantum Gates

In classical computers gates are a small set of circuit elements that are used to implement the combination of binary variables 0's and 1's. Most commonly known gates are AND gate, OR gate and NOT gate.

A ***quantum gate***, a counterpart of classical gate, is a very simple computing device that performs quantum operation on qubits. Quantum gates are one of the essential parts of a quantum computer and are the building blocks of all quantum algorithms.

Quantum gates are mathematically represented as transformation matrices that are *unitary* and the operations performed by these gates are *reversible*. Each unitary transformation U has inverse transformation $U^\dagger$ so that

$$UU^\dagger = I$$
$$U^\dagger = \frac{I}{U} = \frac{1}{U} = U^{-1}$$

Now, the basic question is that *why quantum gates shall be unitary in nature*? It can be explained as follows
A fundamental property of qubits is that they are restricted by the normalization condition, i.e. sum of amplitudes square is equal 1.

$$ie, |\propto|^2 + |\beta|^2 = 1$$

Quantum gates operate on set of qubits and transform them to another quantum state. These operations must preserve the normalization throughout the whole process. The only possible operation for this purpose is unitary matrices. Hence the quantum gates are inevitably *unitary*

Another important feature of quantum gate is that they are *always reversible*. The outputs can be calculated from inputs and inputs can be retrieved from outputs. This is because all unitary matrices are reversible as explained earlier

Note:

1. If the product of a number and its reciprocal is equal to 1, then the number is reversible. For ex

$$2 \times \frac{1}{2} = 1$$

There are different types of quantum gates. *Single-qubit gates* can flip a qubit from 0 to 1 as well as allowing superposition states to be created. *Two-qubit gates* allow the qubits to interact with each other and can be used to create **quantum entanglement** (a strange phenomenon that can't be explained by classical physics).

Some of the important single qubit gates are discussed here. They all are represented by 2 × 2 matrix. (Note that X, Y and Z gates are already discussed earlier under the heading Pauli's matrices. So, it is a sort of repetition)

Single qubit gates

1. X – Gate

This is also called as Pauli X – gate. It is given by

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

When X operates on $|0\rangle$ and $|1\rangle$ the output will be inverted (ie, $|0\rangle$ becomes $|1\rangle$ and $|1\rangle$ becomes$|0\rangle$)

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle$$

$$X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$$

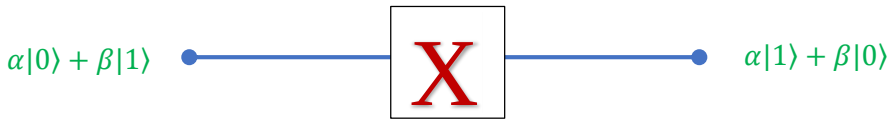
Since X inverts each input it is also called as *bit-flip* gate. If a superposed qubit goes through X gate, the result will be

$$X|\Psi\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \alpha|1\rangle + \beta|0\rangle$$

So,

$$X|\Psi\rangle = \alpha|1\rangle + \beta|0\rangle$$

Symbolically these gates are represented as follows



2. Y – Gate

This is also called as Pauli Y – gate. It is given by

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

When Y operates on $|0\rangle$ and $|1\rangle$

$$Y|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0\times 1 + (-i)\times 0 \\ i\times 1 + 0\times 0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i\begin{bmatrix} 0 \\ 1 \end{bmatrix} = i|1\rangle$$

$$Y|1\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0\times 0 + (-i)\times 1 \\ i\times 0 + 0\times 1 \end{bmatrix} = \begin{bmatrix} -i \\ 0 \end{bmatrix} = -i\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -i|0\rangle$$

So,

$$Y|0\rangle = i|1\rangle \text{ and } \quad Y|1\rangle = -i|0\rangle$$

If a superposed qubit goes through Y gate, the result will be

$$Y|\Psi\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 0\times \alpha + (-i)\times \beta \\ i\times \alpha + 0\times \beta \end{bmatrix} = \begin{bmatrix} -i\beta \\ i\alpha \end{bmatrix} = -i\beta|0\rangle + i\alpha|1\rangle$$

So,

$$Y|\Psi\rangle = i\alpha|1\rangle - i\beta|0\rangle$$

Symbolically these gates are represented as follows

3. Z – Gate

This is also called as Pauli Z – gate. It is given by

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

When Z operates on $|0\rangle$ and $|1\rangle$ the phase will change. Hence this is also called as *phase-flip* gate

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1\times 1 + 0\times 0 \\ 0\times 1 + (-1)\times 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |0\rangle$$

$$Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1\times 0 + 0\times 1 \\ 0\times 0 + (-1)\times 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -|1\rangle$$

So,

$$Z|0\rangle = |0\rangle \text{ and } Z|1\rangle = -|1\rangle$$

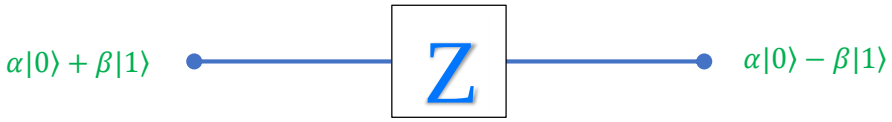
If a superposed qubit goes through Z gate, the result will be

$$Z|\Psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1\times \alpha + 0\times \beta \\ 0\times \alpha + (-1)\times \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix} = \alpha|0\rangle - \beta|1\rangle$$

So,

$$Z|\Psi\rangle = \alpha|0\rangle - \beta|1\rangle$$

Symbolically these gates are represented as follows

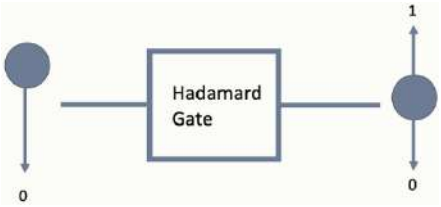


The truth tables for X, Y and Z gates are as follows

X - gate		Y - gate		Z - gate	
Input	**Output**	**Input**	**Output**	**Input**	**Output**
$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$i\|1\rangle$	$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$-i\|0\rangle$	$\|1\rangle$	$-\|1\rangle$
$\alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|1\rangle + \beta\|0\rangle$	$\alpha\|0\rangle + \beta\|1\rangle$	$i\alpha\|1\rangle - i\beta\|0\rangle$	$\alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|0\rangle - \beta\|1\rangle$

4. Hadamard Gate – The gate to superposition

The Hadamard Gate is a well-known gate that brings a qubit into a superposition state. Similar to the Pauli-X gate, the Hadamard Gate acts on a single qubit, and can be represented by a 2 x 2 matrix as follows



Hadamard gate brings a qubit in superposition

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Let us find out what happens when Hadamard gate operates on a qubit that is in the $|0\rangle$ state.

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \times 1 + 1 \times 0 \\ 1 \times 1 + -1 \times 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \quad ---(1)$$

Let us find out what happens when Hadamard gate operates on a qubit that is in the $|1\rangle$ state.

$$H|1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \times 0 + 1 \times 1 \\ 1 \times 0 + -1 \times 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

$$H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} ---(2)$$

If a superposed qubit goes through H gate, the result will be

$$H|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \times \alpha + 1 \times \beta \\ 1 \times \alpha + -1 \times \beta \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \alpha + \beta \\ \alpha - \beta \end{bmatrix}$$

$$H|\psi\rangle = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$$

$$H|\psi\rangle = \alpha\frac{|0\rangle+|1\rangle}{\sqrt{2}} + \beta\frac{|0\rangle-|1\rangle}{\sqrt{2}} - - - (3)$$

The above equations shows that, after applying the Hadamard gate to a qubit that are in $|0\rangle$ & $|1\rangle$ states enter a new superposed states. This is the major difference between X, Y, Z and H gates. In X, Y and Z gates we get only single state whereas in H gate we get superposed state.

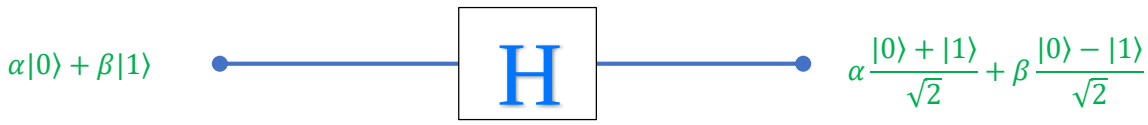
The probability of measuring 0 and 1 is

$$\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}$$

The truth table is as follows

INPUT	OUTPUT
$\|0\rangle$	$\frac{\|0\rangle+\|1\rangle}{\sqrt{2}}$
$\|1\rangle$	$\frac{\|0\rangle-\|1\rangle}{\sqrt{2}}$
$\|\psi\rangle = \alpha\|0\rangle + \beta\|1\rangle$	$\alpha\frac{\|0\rangle+\|1\rangle}{\sqrt{2}} + \beta\frac{\|0\rangle-\|1\rangle}{\sqrt{2}}$

The circuit symbol is as follows



5. Phase Gate (S Gate)

The Phase gate or S gate is a gate that transfers $|0\rangle$ into $|0\rangle$ and $|1\rangle$ into $i|1\rangle$. It is represented as

$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$$

If we apply S gate to a state $|0\rangle$ it will remain same

$$S|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1(1)+0(0) \\ 0(1)+i(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$S|0\rangle = |0\rangle$$

If we apply S gate to a state $|1\rangle$ it will be transformed into $i|1\rangle$

$$S|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1(0)+0(1) \\ 0(0)+i(1) \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$S|1\rangle = i|1\rangle$$

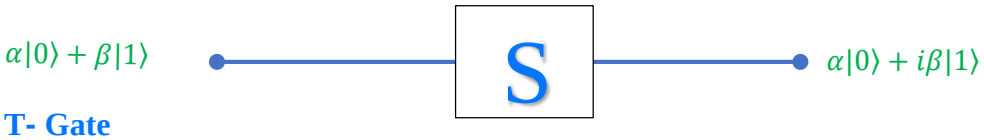
It transforms the state $\alpha|0\rangle + \beta|1\rangle$to the state $\alpha|0\rangle + i\beta|1\rangle$

$$S|\psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ i\beta \end{bmatrix}$$

The truth table is as follows

Input	Output
$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	$i\|1\rangle$
$\|\psi\rangle = \alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|0\rangle + i\beta\|1\rangle$

The symbol is as follows



6. T- Gate

The T-gate is a very commonly used gate and it is given by

$$T = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix}$$

If the input is $|0\rangle$ then the output is also $|0\rangle$

$$T|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$T|0\rangle = |0\rangle$$

If the input is $|1\rangle$ then the output state is $e^{\frac{i\pi}{4}}|1\rangle$

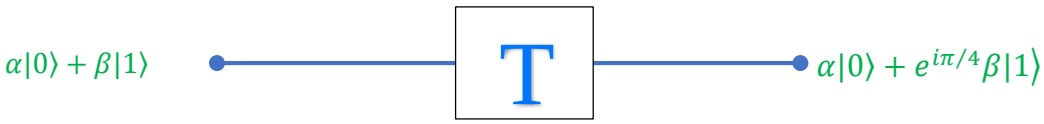
$$T|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ e^{(\frac{i\pi}{4})} \end{bmatrix} = e^{\frac{i\pi}{4}} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$T|1\rangle = e^{\frac{i\pi}{4}}|1\rangle$$

It transforms the state $\alpha|0\rangle + \beta|1\rangle$ to $\alpha|0\rangle + e^{i\pi/4}\beta|1\rangle$

$$T|\psi\rangle = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\pi}{4}} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta e^{\frac{i\pi}{4}} \end{bmatrix}$$

The following figure shows quantum T- gate and the table gives the truth table.



The truth table is as follows

Input	Output
$\|0\rangle$	$\|0\rangle$
$\|1\rangle$	$e^{\frac{i\pi}{4}}\|1\rangle$
$\alpha\|0\rangle + \beta\|1\rangle$	$\alpha\|0\rangle + e^{i\pi/4}\beta\|1\rangle$

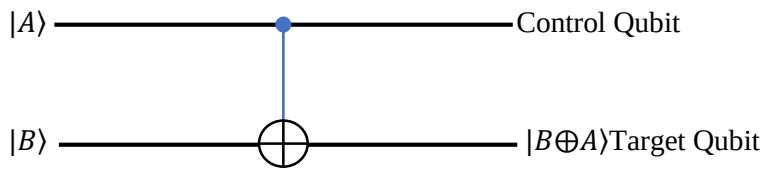
Multiple Qubit gates

As mentioned in the earlier section, Single qubits are interesting, but individually they offer less computational advantage. It is hence essential to look for multiple qubit system and the operation on them. Quantum gates operating on multiple qubits are called as *multiple qubit gates*. Some of them are as follows

1. Controlled Gate (CNOT)

The CNOT gate is a two-qubit operation, where the first qubit is referred as the ***control qubit*** (A) and the second qubit as the ***target qubit*** *(B)*. If the control qubit is $|1\rangle$ then it will flip the target qubit state from$|0\rangle$ to $|1\rangle$ or from $|1\rangle$ to $|0\rangle$. When the control qubit is in state $|0\rangle$ then the target qubit remains unchanged. In fact CNOT applies X on target whenever its control is in state $|1\rangle$

The symbolic representation is as follows. The upper line represents control qubit and bottom line represents target qubit



In the combined qubit, first term is control qubit and the second term is target qubit. For ex, in $|AB\rangle$, A is control qubit and B is target qubit

NOTE: In diagram the control qubit is represented by ● and target is represented by ⊕

Discussion for 4 different input states

1. Input state $|00\rangle$ (Control qubit = 0, Target qubit = 0): Both the bits remain unaltered. Hence, the output state is the same as the input state or $|00\rangle \rightarrow |00\rangle$
2. Input state $|01\rangle$ (Control qubit = 0, Target qubit = 1): Both the bits remain unaltered. Again, the output state is the same as the input state or $|01\rangle \rightarrow |01\rangle$
3. Input state $|10\rangle$ (Control qubit =1, Target qubit = 0): The target qubit is flipped to 1. Therefore, the output state has both qubits 1 or $|10\rangle \rightarrow |11\rangle$
4. Input state $|11\rangle$ (Control qubit =1, Target qubit = 1): The target qubit is flipped to 0. Therefore, the output state becomes $|10\rangle$ or $|11\rangle \rightarrow |10\rangle$.

The truth table of a CNOT gate is as follows

Input	Output
$\|00\rangle$	$\|00\rangle$
$\|01\rangle$	$\|01\rangle$
$\|10\rangle$	$\|11\rangle$
$\|11\rangle$	$\|10\rangle$

We know that two qubits can be in any one of four possible states represented as $|00\rangle$ $|01\rangle|10\rangle$ and $|11\rangle$. The matrix form of them are

$$|00\rangle = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \; |01\rangle = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \; |10\rangle = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \; |11\rangle = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

The state qubit is $|\Psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$. When it is operated by CNOT we get CNOT($\alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$) = $\alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|11\rangle + \alpha_{11}|10\rangle$

From this we can construct the matrix form of CNOT gate as follows (it is 4 ×4 matrix)

The $|00\rangle$remains same as $|00\rangle$. Hence the first column is $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$

The $|01\rangle$remains same as $|01\rangle$. Hence the second column is $\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$

The $|10\rangle$changes to $|11\rangle$. Hence the third column changes from $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$ to $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

The $|11\rangle$changes to $|10\rangle$. Hence the fourth column changes from $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$ to $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$

Hence the matrix form of CNOT gate is

$$CNOT = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\\0 & 0 & 1 & 0\end{bmatrix}$$

Ex (1) S.T the $|00\rangle$remains same as $|00\rangle$ when operated by CNOT

$$\text{CNOT}|00\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\\0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\text{CNOT}|00\rangle = \begin{bmatrix}1+0+0+0\\0+0+0+0\\0+0+0+0\\0+0+0+0\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\therefore \text{CNOT}|00\rangle = |00\rangle$$

Ex (2) S.T the $|01\rangle$remains same as $|01\rangle$ when operated by CNOT

$$\text{CNOT}|01\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\text{CNOT}|01\rangle = \begin{bmatrix}1+0+0+0\\ 0+1+0+0\\ 0+0+0+0\\ 0+0+0+0\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\therefore \text{CNOT}|01\rangle = |01\rangle$$

Ex (3) S.T the $|10\rangle$changes to $|11\rangle$ when operated by CNOT

$$\text{CNOT}|10\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$$

$$\text{CNOT}|10\rangle = \begin{bmatrix}0+0+0+0\\ 0+0+0+0\\ 0+0+0+0\\ 0+0+1+0\end{bmatrix} = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

$$\therefore \text{CNOT}|10\rangle = |11\rangle$$

Ex (4) S.T the $|11\rangle$changes to $|10\rangle$ when operated by CNOT

$$\text{CNOT}|11\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

$$\text{CNOT}|11\rangle = \begin{bmatrix}0+0+0+0\\ 0+0+0+0\\ 0+0+0+1\\ 0+0+0+0\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$$

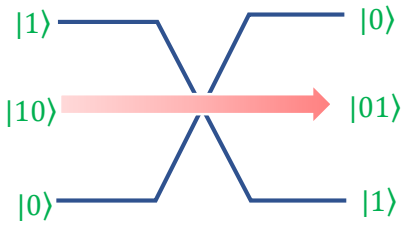
$$\therefore \text{CNOT}|11\rangle = |10\rangle$$

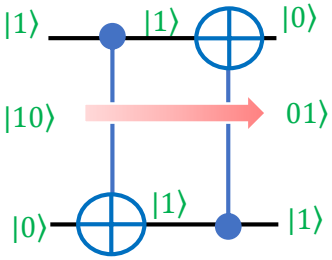
2. Swap Gate

In quantum computation sometimes we need to move state between two qubits, ie from control to target and vice versa. This is nothing but ***swapping*** of the states and the gate used for this purpose is known as SWAP gate.

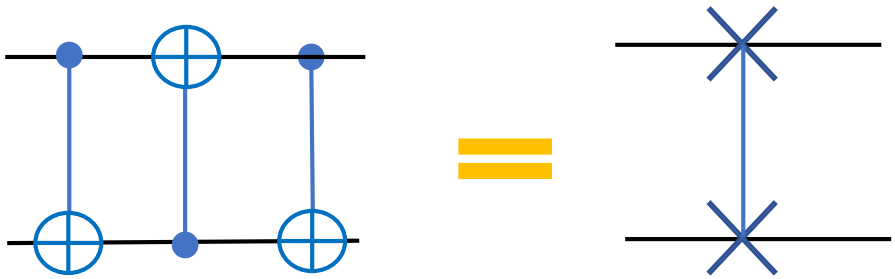
SWAP gate is a two qubit operation gate and swaps the state of the two qubits involved in the operation. It contains 3 CNOT gates.

The action of SWAP gate is explained by taking two CNOT gates as follows where $|10\rangle$ is swapped to $|01\rangle$

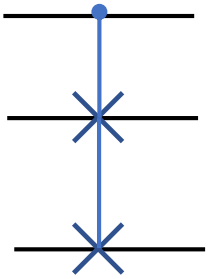




But for effective swapping of the states there must be minimum of 3 CNOT gates. The SWAP circuit is as given below



It is also represented as

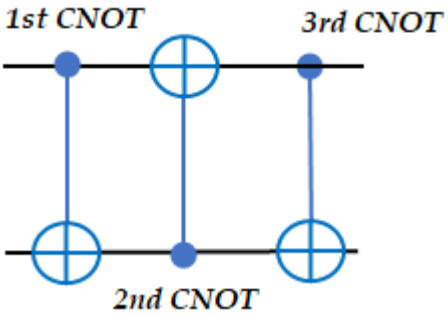


The matrix form of SWAP gate is given by

$$SWAP = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ex (1) S.T the state $|00\rangle$ remains undisturbed by the SWAP gate operation

Consider the SWAP circuit diagram



We know that in CNOT gate if the control qubit is in $|1\rangle$ state then it will flip the target qubit from $|0\rangle$ to $|1\rangle$ and vice versa (otherwise no). So, when $|00\rangle$ is given, the 1st CNOT is **not** satisfied. We stay in the state $|00\rangle$.The 2nd CNOT's control is **not** satisfied. We stay in the state $|00\rangle$.The 3rd CNOT' is also **not** satisfied. We finally stay in the state $|00\rangle$. The same can be verified using matrix analysis as follows

$$SWAP|00\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\end{bmatrix} = \begin{bmatrix}1+0+0+0\\0+0+0+0\\0+0+0+0\\0+0+0+0\end{bmatrix} = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\therefore SWAP|00\rangle = |00\rangle$$

Ex (2) S.T the state $|10\rangle$ is swapped to $|01\rangle$ by SWAP gate operation

$$SWAP|10\rangle = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}0\\0\\1\\0\end{bmatrix} = \begin{bmatrix}0+0+0+0\\0+0+1+0\\0+0+0+0\\0+0+0+0\end{bmatrix} = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\therefore SWAP|10\rangle = |01\rangle$$

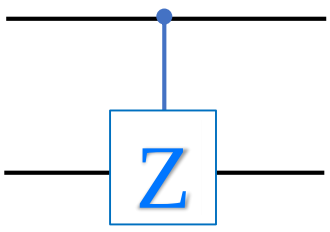
Truth table of swap gate is as follows

Input	Output
$\|00\rangle$	$\|00\rangle$
$\|01\rangle$	$\|10\rangle$
$\|10\rangle$	$\|01\rangle$
$\|11\rangle$	$\|11\rangle$

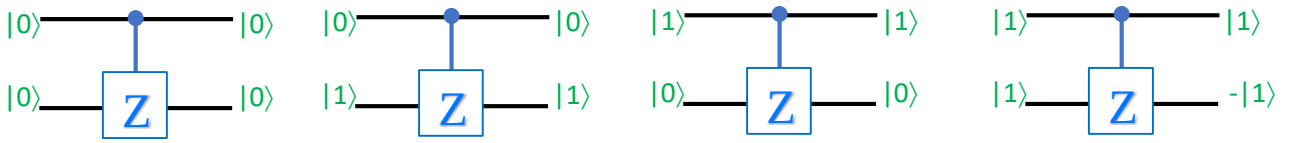
3. Controlled-Z Gate

CNOT gate can be extended in a way that it can work on two qubits based upon a single control qubit. C-Z gate is one such gate. In this gate there is one control qubit and Z unitary matrix as target qubit. If the control qubit is in state $|1\rangle$ then it acts on target Z and will flip the state (ie, there is 180^0 phase change)

The circuit is represented as follows.



Some of the examples are given below



No change because control bit is |0⟩ | *No change because control bit is |0⟩* | *Control bit acts on target but there is no flip of |0⟩* | *Control bit acts on target and flip |1⟩ to -|1⟩*

The truth table of a controlled-Z gate:

Input	Output
|00⟩	|00⟩
|01⟩	|01⟩
|10⟩	|10⟩
|11⟩	-|11⟩

The action of a controlled-Z gate is specified as follows

$$U_Z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Ex (1): S.T the state |10⟩ remains un affected when operated by C-Z gate

$$U_Z|10\rangle = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0+0+0+0 \\ 0+0+0+0 \\ 0+0+1+0 \\ 0+0+0+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\therefore U_Z|10\rangle = |10\rangle$$

Ex (2): S.T the state |11⟩ flips to - |11⟩ when operated by C-Z gate

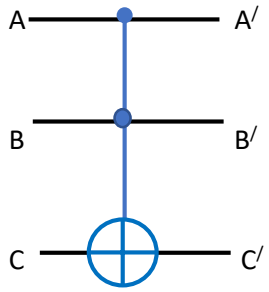
$$U_Z|11\rangle = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0+0+0+0 \\ 0+0+0+0 \\ 0+0+0+0 \\ 0+0+0-1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\therefore U_Z|11\rangle = -|11\rangle$$

4. Toffoli Gate

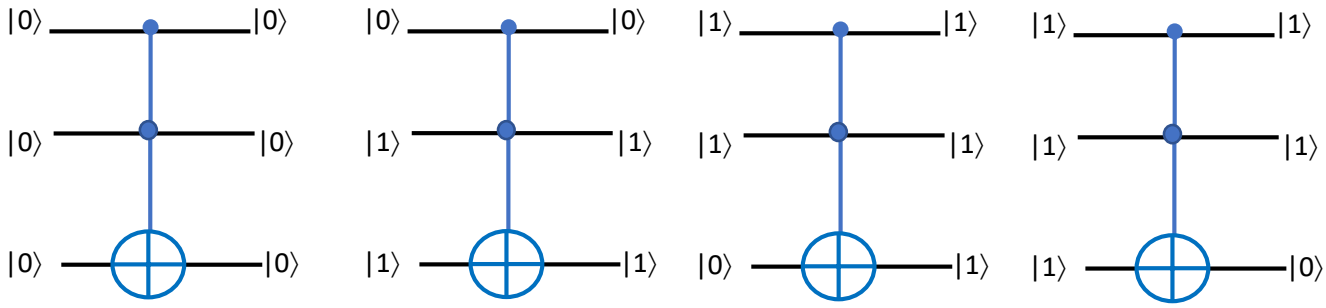
The **Toffoli gate** or **controlled-controlled-NOT (CCNOT)** gate is a logic gate having three input qubits. The first two bits are control bits which remain unaffected by the action of Toffoli Gate. The third is the target bit which is inverted (ie, changes from 0 to 1 or 1 to 0) if both the control bits are 1; else it does not change.

The circuit and the truth table are as follows



Input			Output		
A	B	C	$A^{/}$	$B^{/}$	$C^{/}$
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	0	1	1
1	0	0	1	0	0
1	0	1	1	0	1
1	1	0	1	1	1
1	1	1	1	1	0

Some examples are given here



The Toffoli gate can be expressed as an 8 by 8 matrix as follows

$$U_T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

.

NOTE:

- This is a *reversible* (no information is lost) and *universal* (all reversible logic circuits can be built using Toffoli gates).
- It can be verified that this matrix is unitary and thus the Toffoli gate is a legitimate quantum gate. The quantum Toffoli gate can be used to simulate irreversible classical logic gates and ensures that the quantum gates are capable of performing any computation that a classical computer can do

Limitations of quantum computing

As of now there are some technical difficulties and limitations in building quantum computers. Some of them are

- As the number of quantum gates in a network increases, more interacting qubits are involved, and it is very difficult to monitor their interactions

- The surrounding environment will affect the interactions of qubits (both superposition and entanglement). As a result the quantum information will spread outside the quantum computer and be lost into the environment, thus spoiling the computation. This process is called ***de-coherence*.** How long quantum information will survive before it is spread out is known as ***de-coherency time***
- The number of operations that can be performed before the information is lost due to de-coherency is therefore limited.
- Quantum chips must be kept at very low temperature to create super positions and entanglement of qubits
- The final output of the quantum computers is in the form of a probability. When the question is repeated, the answer changes. Hence repeated operations are required to get correct answer.

Some physicists are pessimistic about the prospects of substantial further progress in quantum computer technology. Some optimistic researchers believe that practical quantum computers will appear in a matter of years rather than decades. We tend towards the optimistic end because

* * * * *

6.4 Question Bank and Numerical Problems

6.4.1 Principles of Quantum Information and Quantum Computing

1. Describe briefly the Quantum Computing and its importance.
2. Elucidate the differences between classical and quantum computing.
3. Define a bit and qubit and explain the differences between them. Mention the properties of Qubit.
4. Discuss the representation of qubit by Bloch Sphere.
5. Explain single, two qubits and extension to N qubits.

6.4.2 Dirac Representation and Matrix Operations

1. Mention the matrix representation of $|0\rangle$ and $|1\rangle$ states and apply the Identity operator to show there is no change in states.
2. State the Pauli matrices and apply Pauli matrices on the states $|0\rangle$ and $|1\rangle$.
3. Explain conjugate and Transpose of a matrix.
4. Describe unitary matrix and along with an example.
5. Describe row and column matrices and the inner product.
6. Discuss probability and quantum superposition.
7. Explain the conditions for orthogonality and orthonormality.

6.4.3 Quantum Gates

1. Illustrate the principle and working of Quantum Not Gate.
2. Discuss the Pauli X,Y and Z gates and their operations on quantum states.
3. Describe the Phase gate along with matrix representation and truth table.
4. Discuss the CNOT gate and its operation on four different input states.
5. explain the matrix form and operation of Toffoli gate.
6. Describe the Swap gate with the matrix and truth table.
7. Elucidate the working of controlled-Z gate mentioning its matrix representation and truth table.
8. Distinguish between single qubit and multiple qubit gates.
9. Describe Quantum Gates with three examples.
10. Discuss the functioning of Hadamard gate with matrix representation and truth table.
11. Explain the working of T gate with its matrix and truth table.

6.5 Numerical Problems

6.5.1 Dirac Representation and Matrix Operations

1. A Linear Operator 'X' operates such that $X|0\rangle = |1\rangle$ and $X|1\rangle = |0\rangle$. Find the matrix representation of 'X'.
2. Given $A = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, Prove that $A^\dagger = A$.
3. Show that the Matrix $U = \begin{bmatrix} \frac{1}{\sqrt{(2)}} & \frac{1}{\sqrt{(2)}} \\ \frac{i}{\sqrt{(2)}} & \frac{-i}{\sqrt{(2)}} \end{bmatrix}$ is Unitary.
4. Find the inner product of states $|1\rangle \, and \, |0\rangle$ and draw conclusions on the result.
5. Given $|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ and $|\phi\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ Prove that $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$

6.5.2 Quantum Gates

1. Using Matrix multiplication show that on applying Hadamard gate twice to a $|0\rangle$ results in its original state.
2. Using two X-gates in series show that two not gates in series are equivalent to a quantum wire.
3. Show the Hadamard Gate is Unitary.
4. Two Qubits are passed through CNOT gate.If the first qubit is the control qubit then what is the output for the following initial states 1. $|00\rangle$, 2. $|01\rangle$, and 3.$|11\rangle$.
5. Show that S gate can be formed by connecting two T gates in Series.

MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2022-23

Module – 4 :

Electrical Properties of Materials and Applications

Notes

Syllabus:

Electrical Properties of Materials and Applications **8 hrs**

Electrical Conductivity in metals

Resistivity and Mobility, Concept of Phonon, Matheissen's rule, Failures of Classical Free Electron Theory, Assumptions of Quantum Free Electron Theory, Fermi Energy, Density of States, Fermi Factor, Variation of Fermi Factor with Temperature and Energy. Numerical Problems.

Superconductivity

Introduction to Super Conductors, Temperature dependence of resistivity, Meissner's Effect, Critical Field, Temperature dependence of Critical field, Types of Super Conductors, BCS theory (Qualitative), Quantum Tunnelling, High Temperature superconductivity, Josephson Junctions (Qualitative), DC and RF SQUIDs (Qualitative), Applications in Quantum Computing: Charge, Phase and Flux qubits, Numerical Problems.

Pre requisites: Basics of Electrical conductivity

Self-learning: Resistivity and Mobility

Electrical Conductivity in Metals

Introduction:

Materials can be classified into three types based on the conductivity of heat and electricity. They are;

1. Conductors (Example : Metals – Copper, Aluminum, Silver, Gold)
2. Semiconductors (Example : Germanium, Silicon)
3. Insulators (Example : Wood, Mica, Glass)

Electron Theory of Metals

The electron theory of metals explains the following:

- Structural, electrical and thermal properties of materials.
- Elasticity, cohesive force and binding in solids.
- Behaviour of conductors, semiconductors, insulators etc.,

In solids, electrons in the outermost orbit of atoms are called valance electrons, which determine the properties of the materials. The electron theory is applicable to all solids (both metals and non-metals). This theory explains the electrical, thermal and magnetic properties of solids.

Classical Free Electron Theory (Drude-Lorentz Theory):

Mobility, Current density, Conductivity and Resistivity.

Mobility of electrons: Mobility is defined as the magnitude of the drift velocity acquired by the electrons in unit electric field. The expression for the mobility is

$$\mu = \frac{v_d}{E} = \frac{e}{m}\tau$$

Current density: Consider a conductor carrying electric current I with area of cross section A perpendicular to the current. The current density J is defined as the ration of current I to the area of cross section A. Hence

$$J = \frac{I}{A}$$

It is observed that the current density J is proportional to the applied electric field E in a conductor. And hence

$$J \propto E \quad \text{or} \quad J = \sigma E$$

The constant of proportionality f is called Electrical Conductivity.

Electrical conductivity in metals: Electrical conductivity of a metal is the ability of the metal to allow electrons to flow through it. The expression for electrical conductivity of a metal is given by

$$\sigma = \frac{ne^2}{m}\tau$$

Where n is the number of free electrons/unit volume & τ is the relaxation time.

Concept of Phonon: A Phonon is a quantum of lattice vibration, the collective motion of atoms constituting a crystal. The Energies and Momenta of Phonons are quantized. It is often characterized as Heat Energy. The study of phonon is an important part of solid state physics. The phonon plays an important role in many of the physical properties of solids such as the thermal conductivity and the electrical conductivity. The conduction electrons in a metal collide against lattice ions during the motion. The interaction is considered to be of type phonon exchange. This results in non-radioactive transitions.

Electrical resistivity: It is the property of the metal and defined as the reciprocal of electrical conductivity.

$$\text{i.e., } \rho = \frac{1}{\sigma} = \frac{m}{ne^2\tau}$$

In metals, the resistivity is due to the scattering of conduction electrons. Scattering of electrons may be due to lattice vibrations or due to impurities.

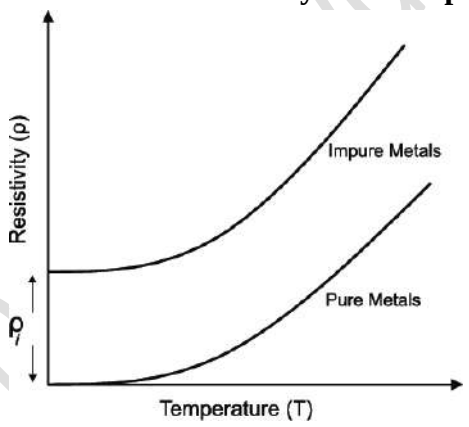
a) Scattering due to lattice vibrations (Phonons) – When the temperature of metal is increased, due to vibrations of lattice ions, the scattering of electrons may take place. The resistivity due to this type of scattering is called 'ideal resistivity' denoted by ρ_{ph}. This is temperature dependent.

b) Scattering due to impurities – The impurities in metals may also scatter electrons and the resistivity due to this is denoted by ρ_i. This is temperature independent.
Thus in any metal, the total resistivity is

$$\rho = \rho_{ph} + \rho_i$$

This is known as Mattthienssen's rule

Variation of resistivity with temperature:



Success of Classical Free Electron Theory:

The successes of this theory are:-

1. It verifies Ohm's law i.e., V = IR
2. It derives Wiedemann-Frenz i.e., $\frac{\sigma_T}{\sigma_E} \propto T$
3. It explains Electrical (σ_E) and Thermal conductivity (σ_T) of metals
4. It explains optical properties of metals.

Failures/Drawbacks of Classical Free-Electron Theory:

1. **Specific Heat (C_V) :** The specific heat of a gas at constant volume is given by
   $C_V = \frac{3}{2}R$, where R is the universal gas constant
   But experimentally it was observed that the specific heat of a metal by its conduction electrons is given by
   $C_V = 10^{-4}RT$
   Thus, the experimental value of C_v is very much lesser than the expected value of C_v. According to classical free electron theory C_v is independent of temperature, but the experimental value of C_v is directly proportional to temperature**.** Hence classical free electron theory fails to explain C_v.
2. **Mean Free Path (λ) :** According to classical free electron theory, the mean free path .
   $\lambda = \bar{c}\tau$, where $\bar{c}$ is the root mean square velocity & τ is the relaxation time of conduction electrons, we also know that $\rho = \frac{m}{ne^2\tau}$ *or* $\tau = \frac{m}{ne^2\rho}$
   By substituting the values of e, m, n & ρ for a metal, τ can be calculated. i.e., $\tau = 2.47\text{x}10^{-14}$ s, where $\bar{c}$ is average velocity of the electron and is equal to $1.15\text{x}10^5$ m/s. In this case, $\lambda = 1.15\text{x}10^5\text{x}2.47\text{x}10^{-14} = 2.85\text{x}10^{-9}$ m = 2.85 nm.
   But the experimental value of λ is found to be 0.285 nm, which is 10 times less than the value obtained from classical free electron theory. Hence classical free electron theory fails to explain λ.
3. **Temperature dependence of electrical conductivity:** According to classical free electron theory, the electrical conductivity of metals is given by
   $$\sigma = \frac{ne^2\tau}{m} \quad or \quad \sigma \propto \tau \quad \rightarrow \quad (1)$$
   On the basis of classical free electron theory, the energy of an electron is given by
   $$\frac{1}{2}mv_{th}^2 = \frac{3}{2}kT \quad \therefore \quad v_{th} = \sqrt{\frac{3kT}{m}} \quad or \quad v_{th} \propto \sqrt{T} \quad \rightarrow \quad (2)$$
   From definition, mean collision time is inversely proportional to v_{th} .
   $$\therefore \ \tau \propto \frac{1}{v_{th}} \quad or \quad \tau \propto \frac{1}{\sqrt{T}} \quad (\because \ v_{th} \propto \sqrt{T}\ ) \quad \rightarrow \quad (3)$$
   From equations (1) & (3), $\sigma \propto \frac{1}{\sqrt{T}}$ $\rightarrow$ (4)
   But experimentally it has been observed that, $\sigma \propto \frac{1}{T}$ $\rightarrow$ (5)
   Hence from equations (4) & (5), it is clear that the prediction of classical free electron theory is not in agreement with the experimental observations. Thus the classical free electron theory fails to explain dependence of T on σ.
4. **Dependence of electrical conductivity on electron concentration:** As per the classical free electron theory, the electrical conductivity of metals is given by
   $$\sigma = \frac{ne^2\tau}{m} \quad or \quad \sigma \propto n$$

Where 'n' is the electron concentration (free electrons)

Hence, divalent & monovalent metals should possess much higher electrical conductivity than monovalent metals. This is contrary to the experimental observations that the monovalent element metals such as copper & silver are more conducting than Zinc (divalent) & aluminum (trivalent).

Or in other words, the theory predicts the direct dependence of electrical conductivity (σ) on number of free electrons per unit volume (n) called number density. But experiments have revealed that $\sigma_{Cu} > \sigma_{Zn} > \sigma_{Al}$ even though the number densities $n_{Cu} < n_{Zn} < n_{Al}$. Hence it fails to explain the dependence of electrical conductivity σ on the number free electrons per unit volume n. Thus the prediction of classical free electron theory that $\sigma \propto n$ does not always hold good. Hence classical free electron theory fails to explain dependence of n on σ.

Concentration of Electrons (n) & Electrical Conductivity (σ) of some metals

Metal	Valency	Electron Concentration n in per m^3	Electrical Conductivity σ in Siemen/meter (S/m)
Cu - 29	**1**	**8.45×10^{28}**	**5.88×10^{7}**
Zn - 30	**2**	**13.10×10^{28}**	**1.09×10^{7}**
Al - 13	**3**	**18.06×10^{28}**	**3.65×10^{7}**

Quantum Free Electron Theory (Summerfeld Theory):

To overcome the drawbacks of classical free electron theory, Sommerfeld proposed quantum free electron theory. He treated electron as a quantum particle. He retains the vital features of classical free electron theory and included the Pauli Exclusion Principle & Fermi-Dirac statistics. The following are the **assumptions** of quantum free electron theory.

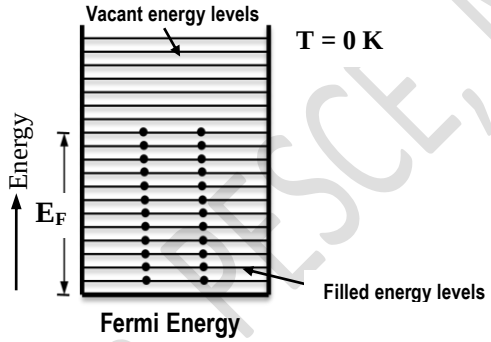
1. The free electrons in a metal can have only discrete energy values. Thus the energies are quantized.
2. The electrons obey Pauli's Exclusion Principle, which states that there cannot be more than two electrons in any energy level.
3. The distribution of electrons in various energy levels obeys the Fermi-Dirac quantum statistics.
4. Free electrons have the same potential energy everywhere within the metal, because the potential due to ionic cores is uniform throughout the metal.
5. The force of attraction between electrons & lattice ions and the force of repulsion between electrons can be neglected.
6. Electrons are treated as wave-like particles.

Fermi - level, Fermi - energy and Fermi - factor

As we know that for a metal containing N atoms, there will be N number of energy levels in each band. According to Pauli's exclusion principle, each energy level can accommodate a maximum of two electrons, one with spin up (+½) and the other with spin down (-½). At absolute zero temperature, two electrons with opposite spins will occupy the lowest available energy level. The next two electrons with opposite spins will occupy the next energy level and so on. Thus, the top most energy level occupied by electrons at absolute zero temperature is called **Fermi-energy level**. The energy corresponding to that energy level is called **Fermi-energy**.

The energy of the highest occupied level at zero-degree absolute is called Fermi energy, and the energy level is referred to as the Fermi level. The Fermi energy is denoted as $\mathbf{E_F}$.

All energy levels below Fermi level are completely filled and above which all energy levels are completely empty.



At temperatures above absolute zero, the electrons get thermally excited and move up to higher energy levels. As a result, there will be many vacant energy levels below as well as above Fermi energy level. Under thermal equilibrium, the distribution of electrons among various energy levels is given by statistical function *f*(E). The function ***f*(E)** is called **Fermi-factor** and this gives the probability of occupation of a given energy level under thermal equilibrium. The expression for *f*(E) is given by

$$f(E) = \frac{1}{e^{(E-E_F)/kT} + 1}$$

Where ***f*(E)** is called Fermi-Dirac distribution function of Fermi factor, $\mathbf{E_F}$ is the Fermi energy, **k** is the Boltzmann constant and **T** is the temperature of metal under thermal equilibrium.

Note: 1. The Fermi-Dirac distribution *f*(E) is used to calculate the probability of an electron occupying a certain energy level.
2. The distribution of electrons among the different energy levels as a function of temperature is known as Fermi-Dirac distribution function.

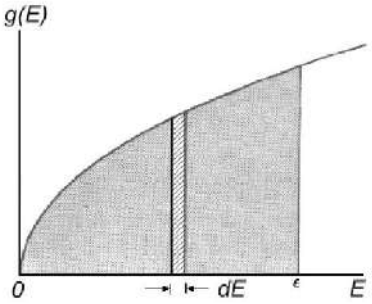
Density of States (DoS) :

According to band theory, energy bands are formed in solids and in a band the spacing between two successive energy levels decreases with increase in energy.

The Density of States is defined as the number of energy states available per unit volume of the material in the unit energy range in the valence band of the material. It is mathematically a continuous function denoted g(E).

The number of energy levels in the energy range E and E+dE per unit volume of the material is given by g(E) dE.

$$g(E)dE = \left(\frac{8\sqrt{2}\pi m^{3/2}}{h^3}\right)E^{1/2}dE$$



Density of states function vs Energy

Variation of Fermi factor with Energy and Temperature

Let us consider the different cases by considering the Fermi factor equation

$$f(E) = \frac{1}{e^{(E-E_F)/kT}+1}$$

Case (i) : *f*(E) for E < E_F at T = 0 K;

When; E < E_F & T = 0 K, from the probability function *f*(E) we have

$$f(E) = \frac{1}{e^{-\infty}+1} = \frac{1}{0+1} = 1$$

i.e., *f*(E) = 1 for E < E_F at T = 0 K.

This implies that at absolute zero temperature, all the energy levels below E_F are 100% occupied which is true from the definition of Fermi energy.

Case (ii) : *f*(E) for E > E_F at T = 0 K;

When E > E_F & T = 0 K, then *f*(E) becomes

$$f(E) = \frac{1}{e^{\infty}+1} = \frac{1}{\infty+1} = \frac{1}{\infty} = 0$$

i.e., *f*(E) = 0 for E > E_F at T = 0 K.

This implies that at absolute zero temperature, all the energy levels above E_F are unoccupied (completely empty) which is true from the definition of Fermi energy.

Case (iii) : *f*(E) for E = E_F at T = 0 K;

When E = E_F & T = 0 K, then *f*(E) becomes

$$f(E) = \frac{1}{e^{\%} + 1} = \text{Indetermin ate}$$

i.e., $f(E) = \infty$ for $E = E_F$ at $T = 0$ K.

Hence, the occupation of Fermi level at T = 0 K has an undetermined value ranging between zero and unity (0 & 1). The Fermi-Dirac distribution function is discontinuous at $E = E_F$ for T = 0 K.

Case (iv) :- $f(E)$ for $E = E_F$ at $T > 0$ K;

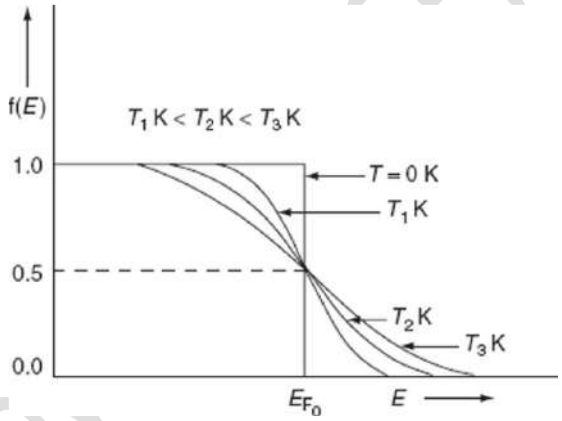
When $E = E_F$ & $T > 0$ K, then $f(E)$ becomes

$$f(E) = \frac{1}{e^{0} + 1} = \frac{1}{1+1} = \frac{1}{2}$$

i.e., $f(E) = ½$ for $E = E_F$ at $T > 0$ K.

If $E \ll E_F$, the probability starts decreasing from 1 and reaches 0.5 (½) at $E = E_F$ and for $E > E_F$, it further falls off as shown in figure. In conclusion, the Fermi energy is the most probable or average energy of the electrons in a solid.

The variation of Fermi factor with energy and temperature is as shown in figure given below.



Importance of Fermi Energy

- Fermi energy level is used to separate the vacant and filled states at 0 K.
- It is used to know the status of the electrons.
- Electrons are completely filled below the Fermi energy level and completely empty above the Fermi level at 0 K.
- Above 0 K some electrons absorb thermal energy and they jump to the higher energy levels.

Fermi Temperature (T_F) : - Fermi temperature is the temperature at which the average thermal energy of the free electron in a solid is equal to the Fermi energy at 0 K.

But the thermal energy possessed by electrons is given by the product kT.

Thus, when $T = T_F$, $kT_F = E_{F_0}$ is satisfied

But all practical purposes, $E_{F_0} = E_F$ $\therefore \quad kT_F = E_F \quad \therefore \quad T_F = \frac{E_F}{k}$

This is the expression for Fermi temperature.

For metals, we know that E_F will be of the order of few eV, say E_F = 5 eV, then

$$T_F = \frac{5x1.602x10^{-19}}{1.38x10^{-23}} = 58000\ K$$

Fermi velocity (v_F) : - The energy of the electrons in metals at Fermi level is E_F. The velocity of the electrons which occupy the Fermi level is called Fermi velocity v_F.

$$E_F = \frac{1}{2}mv_F^2$$

or

$$v_F = \left(\frac{2E_F}{m}\right)^{1/2}$$

This is the expression for Fermi velocity.

Merits or Success of Quantum free electron theory:

The quantum free electron theory solves the flaws of the classical free electron theory which are discussed below.

1. **Specific heat of free electrons:** According to quantum free electron theory, the electrons occupying energy levels close to E_F can absorb heat energy. Such electrons constitute a very small percentage of the total number of free electrons. Hence the specific heat of free electrons is given by

   $$C_V = \frac{2k}{E_F}RT$$

   Since the value of E_F ranges from 1 to 10 eV, by taking a typical value of E_F = 5 eV, we get

   $$\frac{2k}{E_F} \approx 10^{-4}$$

   $$\therefore C_V = 10^{-4}RT$$

which is in agrees well with the experimental results.

2. **Temperature dependence of resistivity or conductivity in metals:** According to quantum free electron theory, the expressions for electrical conductivity & resistivity of a metal are given by

   $$\sigma = \frac{ne^2}{m^*}\left(\frac{\lambda}{v_F}\right) \quad or \quad \rho = \frac{m^*v_F}{ne^2\lambda}\left(\frac{v_F}{\lambda}\right)$$

In the above expression only the mean free path λ is the temperature dependent quantity.

$$\therefore \sigma \propto \lambda \quad \rightarrow \quad (1)$$

In classical theory, the collision was seen as a particle bouncing off another. In the quantum understanding, an electron is viewed as a wave travelling through the medium. If **r** represents the amplitude of the oscillation of the lattice ions can be considered to present a circular cross

section of area πr^2 that blocks the path of the electron waves. Hence electron waves are scattered more effectively results in a reduction of mean free path (λ). Thus λ is inversely proportional to the area of cross section.

$$\text{i.e.,} \quad \lambda \propto \frac{1}{\pi r^2} \quad \rightarrow \quad (2)$$

But the area of cross section πr^2 is directly proportional to the absolute temperature.

$$i.e., \quad \pi r^2 \propto T$$

$$\therefore \quad \lambda \propto \frac{1}{T} \quad \rightarrow \quad (3)$$

By substituti ng the values of λ in equation (1) we get

$$\sigma \propto \frac{1}{T} \qquad \text{or} \qquad \rho \propto T$$

This is exactly same as the experimental prediction. Thus quantum free electron theory properly explains the dependence of σ on T.

3. **Dependence of electrical conductivity on electron concentration:** According to quantum free electro theory, the electrical conductivity in metals is given by

$$\sigma = \frac{ne^2}{m^*}\left(\frac{\lambda}{v_F}\right)$$

From the above equation it is clear that the electrical conductivity depends on both the electron concentration n and $\left(\frac{\lambda}{v_F}\right)$.

If we compare the cases of copper and aluminium, the value of n for aluminium is 2.13 times higher than that of copper. But the value of λ/v_F for copper is about 3.73 times higher than that of aluminium. Thus the conductivity of copper is more than that of aluminium.

Problems :

1. The free electron density of aluminium is 18.10×10^{28} m^{-3}. Calculate its Fermi energy at 0 K. Planck's constant and mass of free electron are 6.626×10^{-34} Js and 9.1×10^{-31} kg.

Solns.

n = 18.10×10^{28} m^{-3}

Planck's constant, h = 6.626×10^{-34} Js

Mass of an electron, m = 9.11×10^{-31} kg

$$E_{F_0} = \left(\frac{h^2}{8m}\right)\left(\frac{3n}{\pi}\right)^{2/3} = \left(\frac{(60626x10^{-34})^2}{8x9.11x10^{-31}}\right)\left(\frac{3x18.10x10^{28}}{\pi}\right)^{2/3}$$

$$= 1.8689x10^{-18}\,J = \frac{1.8689x10^{-18}}{1.602x10^{-19}} = 11.68 \quad eV$$

2. Calculate the density of states for copper at the Fermi level for T = 0 K. Given that, electron density of copper is 8.5×10^{28} electrons /m^3.

Solns.

n = $8.5x10^{28}$ m^{-3}

Planck's constant, h = $6.626x10^{-34}$ Js

Mass of an electron, m = $9.1x10^{-31}$ kg

$$g(E)dE = \frac{\pi}{2}\left(\frac{8m}{h^2}\right)^{3/2} E_F^{\frac{1}{2}} dE \quad where, \quad E_F = \left(\frac{h^2}{8m}\right)\left(\frac{3n}{\pi}\right)^{2/3}$$

$$E_F = \left(\frac{h^2}{8m}\right)\left(\frac{3n}{\pi}\right)^{2/3} = \left(\frac{\left(6.626x10^{-34}\right)^2}{8x9.1x10^{-31}}\right)\left(\frac{3x8.5x10^{28}}{\pi}\right)^{2/3} = 1.1293x10^{-18}\,J = 7.05\ eV$$

$$g(E) = \frac{\pi}{2}\left(\frac{8m}{h^2}\right)^{3/2} E_F^{\frac{1}{2}} = \frac{\pi}{2}\left(\frac{8x9.1x10^{-31}}{\left(6.626x10^{-34}\right)^2}\right)^{3/2} x\ 7.05^{1/2}$$

3. Find the probability of an electron occupying an energy level 0.02 eV above the Fermi level at 200 K and 400 K in a material.

Solns.

E - E_F = 0.02 eV = 0.02x1.602x10^{-19} J = 3.204 x10^{-21} J

T_1 = 200 K & T_2 = 400 K

$$f(E)_{atT_1} = \frac{1}{e^{(E-E_F)/kT_1}+1} == \frac{1}{e^{\left(\frac{3.204x10^{-21}}{1.38x10^{-23}x200}\right)}+1} = \frac{1}{e^{(1.1594)}+1} = \frac{1}{3.188+1} = 0.24$$

$$f(E)_{atT_2} = \frac{1}{e^{(E-E_F)/kT_2}+1} == \frac{1}{e^{\left(\frac{3.204x10^{-21}}{1.38x10^{-23}x400}\right)}+1} = \frac{1}{e^{(0.5797)}+1} = \frac{1}{1.7855+1} = 0.36$$

4. Show that the sum of the probability of occupancy of an energy state at ΔE above the Fermi level and that at ΔE below the Fermi level is unity.

Solns.

Let $f(E)_E^a$ is the probabilit y of ocuupation abobe the Fermi level at energy $E = (E_F + \Delta E)$

&

$f(E)_E^b$ is the probabilit y of ocuupation below the Fermi level at $E = (E_F - \Delta E)$

$S.T.$ $\quad f(E)_E^a + f(E)_E^b = 1$

We know that $f(E) = \frac{1}{e^{(E-E_F)/kT}+1}$

For energy above the Fermi level, $\quad E = (E_F + \Delta E)$ or $E - E_F = \Delta E \quad \therefore \quad f(E)_E^a = \frac{1}{e^{(\Delta E)/kT}+1}$

For energy below the Fermi level, $\quad E = (E_F - \Delta E)$ or $E - E_F = -\Delta E \quad \therefore \quad f(E)_E^b = \frac{1}{e^{-(\Delta E)/kT}+1}$

$$\therefore f(E)_E^a + f(E)_E^b = \frac{1}{e^{(\Delta E)/kT}+1} + \frac{1}{e^{-(\Delta E)/kT}+1}$$

Put $e^{(\Delta E)/kT} = x, \quad \therefore \quad e^{-(\Delta E)/kT} = \frac{1}{x}, \quad$ Hence.,

$$f(E)_E^a + f(E)_E^b = \frac{1}{x+1} + \frac{1}{\frac{1}{x}+1} = \frac{1}{x+1} + \frac{1}{\frac{1+x}{x}} = \frac{1}{x+1} + \frac{x}{1+x} = \frac{1+x}{x+1} = 1$$

$$\therefore \quad f(E)_E^a + f(E)_E^b = 1$$

5. Calculate the probability of an electron occupying an energy level of 0.05 eV at 200 K above and below the Fermi level.

Soln.

$E - E_F = 0.05\,eV, \quad T = 200\,K$

$f(E)_E^a = ? \quad \& \quad f(E)_E^b = ?$

For energy above the Fermi level, $E = (E_F + \Delta E)$ or $E - E_F = 0.05\,eV = 0.05x1.602x10^{-19}\,J$

$$\therefore \text{ At, } T = 200\,K., \quad f(E)_E^a = \frac{1}{e^{\left(\frac{0.05x1.602x10^{-19}}{1.38x10^{-23}x500}\right)}+1} = \frac{1}{e^{1.16087}+1} = \frac{1}{3.1927+1} = \frac{1}{4.1927} = 0.2385$$

For energy below the Fermi level, $E = (E_F - \Delta E)$ or $E - E_F = -\Delta E \quad \therefore \quad f(E)_E^b = \frac{1}{e^{-(\Delta E)/kT}+1}$

$$\therefore \text{ At, } T = 200\,K., \quad f(E)_E^b = \frac{1}{e^{-\left(\frac{0.05x1.602x10^{-19}}{1.38x10^{-23}x500}\right)}+1} = \frac{1}{e^{-1.16087}+1} = \frac{1}{0.3132+1} = \frac{1}{1.3132} = 0.7615$$

$$\therefore f(E)_E^a + f(E)_E^b = 0.2385 + 0.7615 = 1$$

6. Find the temperature at which there is 1 % probability that a state with 0.5 eV energy above the Fermi energy is occupied.

Soln.

$E - E_F = 0.5\,eV = 0.5x1.602x10^{-19}\,J$

$f(E) = 1\% = 0.01$

$T = ?$

$$W.K.T \qquad f(E) = \frac{1}{e^{(E-E_F)/kT}+1} \qquad or \qquad 0.01 = \frac{1}{e^{\left(\frac{0.5x1.602x10^{-19}}{1.38x10^{-23}xT}\right)}+1} = \frac{1}{e^{\left(\frac{5804}{T}\right)}+1}$$

$$or \qquad 0.01 = \frac{1}{e^{\left(\frac{0.5x1.602x10^{-19}}{1.38x10^{-23}xT}\right)}+1} = \frac{1}{e^{\left(\frac{5804}{T}\right)}+1}$$

$$\therefore \qquad e^{\left(\frac{5804}{T}\right)} + 1 = \frac{1}{0.01} = 100 \quad \text{or} \quad e^{\left(\frac{5804}{T}\right)} = 100 - 1 = 99$$

$$\frac{5804}{T} = \ln(99) = 4.595 \quad \text{or} \quad T = \frac{5804}{4.595} = 1263K$$

7. The Fermi level in potassium is 2.1 eV. What are the energies for which the probability of occupancy at 300 K are 0.99, 0.01 and 0.5?

Soln.

For potassium $E_F = 2.1\,eV$

Find, $E_1 = ?$ for $f(E_1) = 0.99$

Find, $E_2 = ?$ for $f(E_2) = 0.01$

& Find, $E_3 = ?$ for $f(E_3) = 0.5$

$$W.K.T \qquad f(E) = \frac{1}{e^{(E-E_F)/kT}+1} \quad \text{or} \quad e^{(E-E_F)/kT} = \frac{1}{f(E)} - 1$$

Taking natural logarithm on both sides, we get,

$$\frac{E-E_F}{kT} = \ln\left(\frac{1}{f(E)} - 1\right) \quad \text{or} \quad E = E_F + kT\ln\left(\frac{1}{f(E)} - 1\right)$$

$\therefore E_1$ for $f(E_1)$ is given by

$$E_1 = E_F + kT\ln\left(\frac{1}{f(E_1)} - 1\right) = 2.1 + \frac{1.38x10^{-23}x300}{1.602x10^{-19}}\ln\left(\frac{1}{0.99} - 1\right)$$

$$= 2.1 + 0.02584x(-4.5950) = 2.1 - 0.1187 = 1.9813\,eV$$

slly E_2 for $f(E_2)$ is given by

$$E_2 = E_F + kT\ln\left(\frac{1}{f(E_2)} - 1\right) = 2.1 + \frac{1.38x10^{-23}x300}{1.602x10^{-19}}\ln\left(\frac{1}{0.01} - 1\right)$$

$$= 2.1 + 0.02584x(4.5950) = 2.1 + 0.1187 = 2.2187\,eV$$

slly E_3 for $f(E_3)$ is given by

$$E_3 = E_F + kT\ln\left(\frac{1}{f(E_3)} - 1\right) = 2.1 + \frac{1.38x10^{-23}x300}{1.602x10^{-19}}\ln\left(\frac{1}{0.5} - 1\right)$$

$$= 2.1 + 0.02584x(0) = 2.1 + 0 = 2.1\,eV$$

8. Calculate the Fermi energy in eV for a metal at 0 K, whose density is 10500 kgm^{-3}, atomic weight is 107.9 and it has one conduction electron per atom.

Solns.

Density of metal, ρ = 10500 kgm^{-3}

Atomic weight of metal, wt. = 107.9

Fermi energy, E_F = ?

We know that the concentration of electrons in metal, n is given by

$$n = \frac{\text{density}(\rho)\ \text{x}\ \text{Avogadro's constant}(N_A)\ \text{x}\ \text{no. of free electrons per atom}}{\text{atomic weight (wt.)}}$$

$$= \frac{10500\ \text{x}\ 60022x10^{26}\ \text{x}\ 1}{107.9} = 5.861x10^{28}m^{-3}$$

$$E_{F_0} = \left(\frac{h^2}{8m}\right)\left(\frac{3}{\pi}\right)^{2/3}(n)^{2/3} = \left(\frac{(60626x10^{-34})^2}{8x9.11x10^{-31}}\right)\left(\frac{3x5.816x10^{28}}{\pi}\right)^{2/3}$$

$$= 8.8173x10^{-18}J = \frac{8.8173x10^{-18}}{1.602x10^{-19}} = 5.51 \quad eV$$

REVIEW QUESTIONS

Long Answer Questions:

1. Define the terms: Fermi-level, Fermi-energy and Fermi-factor.
2. Mention the drawback/Failures of classical free electron theory and explain them.
3. What are the assumptions of quantum free electron theory?
4. Define the terms: Fermi-temperature, Fermi-velocity and density of states.
5. What are the successes of quantum free electron theory? Explain.
6. How quantum free electron theory successfully explain the failures of classical free electron theory.
7. What are the merits of quantum free electron theory? Explain.
8. Write down the Fermi-Dirac equation for the probability of occupation of an energy level E by an electron. Show that the probability of its occupancy by an electron is zero if $E > E_F$ and unity if $E < E_F$ at temperature 0 K.
9. Define Fermi factor. Explain Fermi Dirac distribution for electrons in a metal at temperature T = 0 K and T > 0 K.
10. Write an expression for the Fermi energy distribution function *f*(E) and discuss its behavior with change in temperature. Plot *f*(E) versus E for T = 0 K, and T > 0 K.
11. Explain the dependence of electrical conductivity on temperature and electron concentration on the basis of classical free electron theory.
12. Write down the difference between classical and quantum free electron theories.

Problems:

1. Obtain the value of *f*(E) for $E - E_F = 0.01$ eV at T = 300 K.
2. Find the probability that an energy level at 0.2 eV below Fermi level being occupied at temperatures 300 K and 900 K.
3. At what temperature can we expect a 10% probability that electron in silver have an energy which 1% above the Fermi energy? The Fermi energy of silver is 5.5 eV.
4. Evaluate the Fermi function for an energy 0.04 eV at T= 330 K above the Fermi energy.
5. Show that the probability of occupation above the fermi level is same as the non-occupation probability below the fermi level for given energy and temperature.

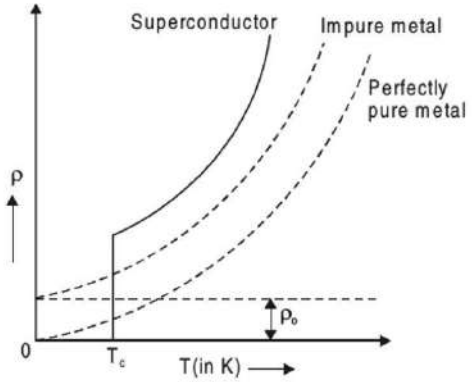
*** * * * *END* * * * ***

Superconductivity:

Introduction: Lord Kamerlingh Onnes discovered the phenomenon of superconductivity in the year 1911. When he was studying the temperature dependence of resistance of Mercury at very low temperature he found that resistance of Mercury decreases in temperature up to a particular temperature T_c = 4.15 K and below this temperature the resistance of mercury abruptly drops to zero. Between 4.15K and 0K Mercury offered no resistance for the flow of electric current. This phenomenon is reversible and material becomes normal when once again temperature was increased above 4.15 K. This phenomenon is called superconductivity and material which exhibits this property is named as superconductor.

Definition: Superconductivity is defined as "The phenomenon in which resistance of certain metals, alloys and compounds drops to zero abruptly, below certain temperature is called superconductivity.

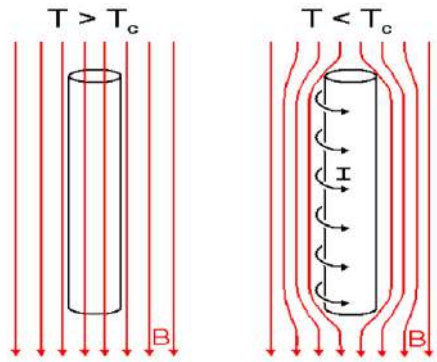
Variation of Resistivity with Temperature: The variation of the resistivity of a superconductor, pure and impure metals with temperature is as shown in the figure below.



Critical Temperature: The temperature, below which materials exhibit superconducting property is called critical temperature, denoted by T_C. It is different for different substances. The materials, which exhibit superconducting property, are called superconductors. Above critical temperature material is said to be in normal state and offers resistance for the flow of electric current. Below critical temperature material is said to be in superconducting state. Thus T_C is also called as transition temperature.

Meissner's Effect

In 1933, Meissner and Ochsenfeld showed that when a superconducting material is placed in a magnetic field it allows magnetic lines of force to pass through, if its temperature is above T_C. If the temperature is reduced below the critical temperature T_C, then it expels all the flux lines completely out of the specimen and exhibits perfect diamagnetism. This is known as Meissner's effect. Since superconductor exhibits perfect diamagnetism below the critical temperature Tc, magnetic flux density inside the material is zero.



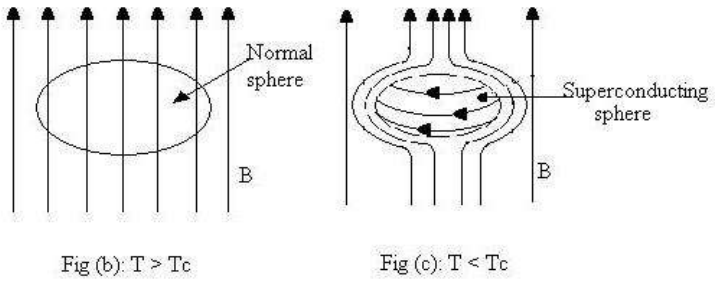


Fig (b): T > Tc Fig (c): T < Tc

The expression for magnetic flux density is given by $B = \mu_0 (M + H)$

Here B is Magnetic Flux Density, M is Magnetization and H is the applied magnetic field strength. For a superconductor, B = 0 at $T < T_C$.

Thus we get M = − H.

Thus Meissner's Effect signifies the negative magnetic moment associated with superconductors.

Critical Field and its Temperature Dependence

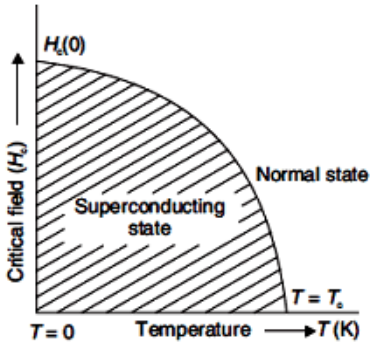
Critical field We know that when superconductor is placed in a magnetic field it expels magnetic flux lines completely out of the body and exhibits a perfect diamagnetism. But if the strength of the magnetic field is further increased, it is found that for a particular value of the magnetic field,

material looses its superconducting property and becomes a normal conductor. The value of the magnetic field at which the transition occurs from the Superconducting state to Normal Conducting state is called Critical Field or Critical Magnetic Field and is denoted by H_C. It is found that by reducing the temperature of the material further superconducting property of the material could be restored. Thus, critical field does not destroy the superconducting property of the material completely but only reduces the critical temperature of the material.

The variation of Critical field with temperature below the critical temperature is given by

$$H_C = H_0\left(1-\frac{T^2}{T_C^2}\right)$$

Here H_C is the Critical field at any temperature T less than T_C, H_0 is the Critical field at T = 0 K.



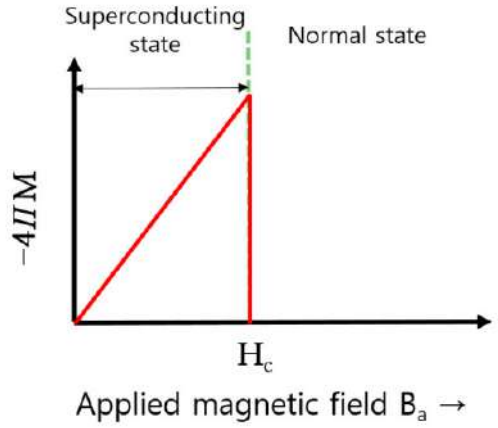
Types of Superconductors

Superconductors are classified into two types

1. Type - I Superconductor or Soft Superconductor
2. Type - II Superconductor or Hard Superconductor

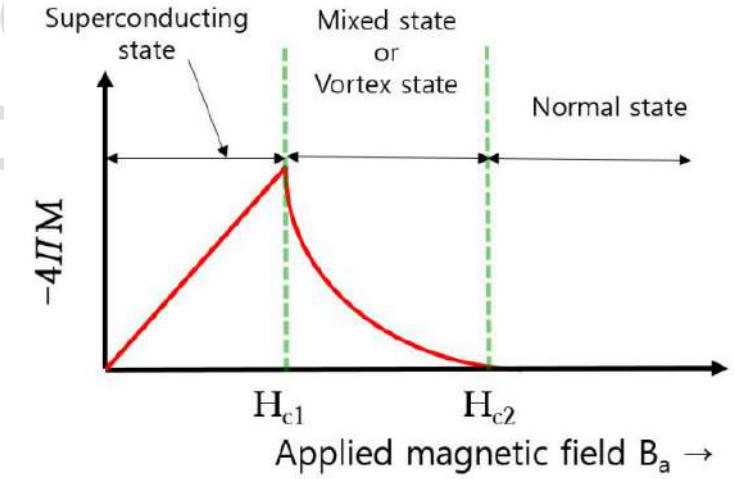
Type - I Superconductors:

Type I superconductors exhibit complete Meissener's Effect and possess a single value of critical field. The graph of magnetic moment Vs magnetic field is as shown in the Fig. below. As the field strength increases the material becomes more and more diamagnetic until H becomes equal to H_C. Above H_C the material allows the flux lines to pass through and exhibits normal conductivity. The value of H_C is very small for soft superconductors. Therefore, soft superconductors cannot withstand high magnetic fields. Therefore, they cannot be used for making superconducting magnets. Ex. Hg, Pb and Zn.



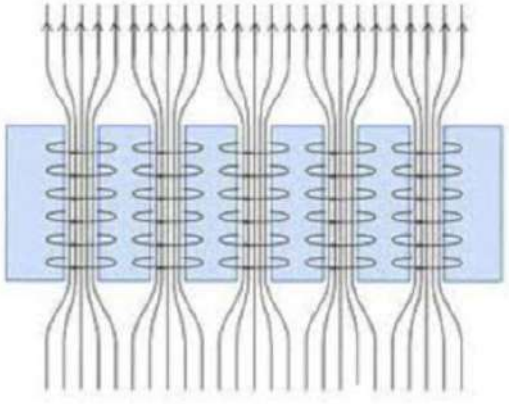
Type - II Superconductors:

paragraph Superconducting materials, which can withstand high value of critical magnetic fields, are called Hard Superconductors.



The graph of magnetic moment Vs magnetic field is as shown in the Fig. Hard superconductors are characterized by two critical fields H_{C1} and H_{C2}. When applied magnetic field is less than H_{C1} material exhibits perfect diamagnetism. Beyond H_{C1} partial flux penetrates and the material is said to be Vortex State. Thus flux penetration occurs through

small-channelized regions called filaments. As the strength of the field increases further, more and more flux fills the body and thereby decreasing the diamagnetic property of the material. At H_{C2} flux fills the body completely and material losses its diamagnetic property as well as superconducting property completely.



The value of H_{C2} is hundreds of times greater than H_C of soft superconductors. Therefore, they are used for making powerful superconducting magnets. Ex. : *NbTi, Nb_3Sn.*

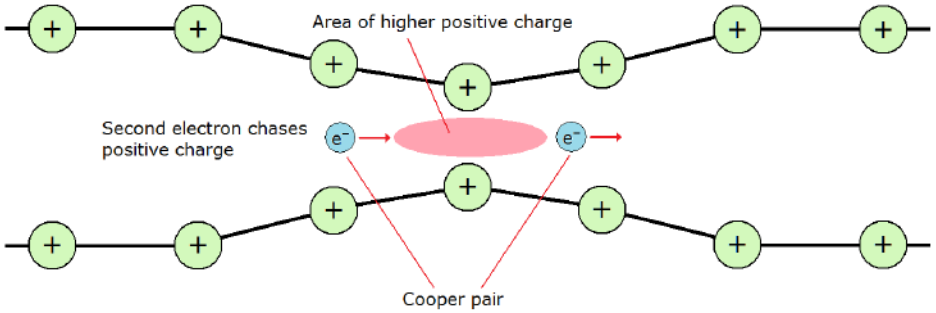
Difference between Type - 1 and Type -2 Superconductors	
Type - 1 Superconductor	**Type -2 Superconductor**
1. They exhibit complete Meissner effect	1. They exhibit partial Meissner effect
2. These are perfect diamagnetics	2. These are not perfect diamagnetics
3. These are known as soft superconductors	3. These are known as hard superconductors
4. They have only one critical magnetic field	4. They have two critical magnetic fields
5. No vertex (Mixed) state is present	5. Vertex (mixed) state is present
6. These materials undergoes a sharp transition at the critical magnetic field	6. These materials undergoes a gradual transition between two critical magnetic fields
7. The highest value of critical magnetic field is 0.1 wb/m^2	7. The upper critical magnetic field is of the 50 wb/m^2
8. Critical temperature is low (< 10 K)	8. Critical temperature is high (> 10 K)
9. Applications are very limited	9. They are used to generate very high magnetic field.
10. Examples:- lead, tin, mercury , etc.	10. Examples:- alloys like Nb-Sn, Nb-Ti, Nb-Zr, etc.

BCS Theory of Superconductivity:

Bardeen, Cooper and Schrieffer explained the phenomenon of superconductivity in the year 1957. The essence of the BCS theory is as follows.

Consider an electron approaching a positive ion core and suffers attractive coulomb interaction. Due to this attraction ion core is set in motion and thus distorts that lattice. Let a second electron come in the way of distorted lattice and interaction between the two occurs which lowers the energy of the second electron. The two electrons therefore interact indirectly through the lattice distortion or the phonon field which lowers the energy of the electrons. The above interaction is interpreted as electron - Lattice – electron interaction through phonon field.

It was shown by Cooper that, this attractive force becomes maximum if two electrons have opposite spins and momentum. The attractive force may exceed coulombs repulsive force between the two electrons below the critical temperature, which results in the formation of bound pair of electrons called cooper pairs.



Below the critical temperature the dense cloud of Cooper pairs form a collective state and the motion all Cooper pairs is correlated resulting in zero resistance of the material.

High Temperature Superconductors:

Superconducting materials which exhibit superconductivity at relatively higher temperatures are called high temperature superconductors. Thus high temperature superconductors possess higher value of critical temperature compared to conventional superconductors. Most of the high temperature superconductors are found to fall into the category of ceramics. In 1986 George Bednorz and Alex Muller discovered a compound containing Lanthanum, Barium, Copper and Oxygen having T_C = 30 K was developed. In 1987 scientists developed a compound which is an oxide of the form $YBa_2Cu_3O_7$ which is referred to as 1-2-3 compound with T_C > 90 K was discovered.

All high temperature superconductors are oxides of copper and bear Perovskite crystal structure characterized by large number of copper-oxygen layers. It was found that addition of extra copper-oxygen layer pushes the critical temperature T_C to higher values. The super currents are strong in the copper-oxygen layer and weak in the direction perpendicular to the planes.

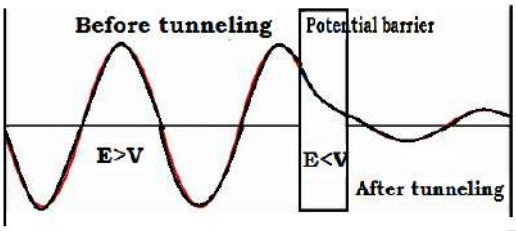
High Temperature superconductors are not commercially available mainly due to their current densities and difficulty in forming into wires. Once these difficulties are overcome such superconductors find many applications in various fields like zero-loss power transmission lines, super- strong magnetic materials and as the materials for levitating trains.

Following is the list of some of High Temperature Superconductors.

Superconductor	Year	T_C (K)
K_yWO_3	1967	6.0
$LiTi_{2+y}O_4$	1973	1.2
$BaPb_{1-y}Bi_yO_3$	1975	13
$La_{2-y}Ba_yCuO_4$	1986	30
$YBa_2Cu_3O_{7-y}$	1987	90
$Ba_{1-y}K_yBiO_3$	1988	20
$BiSrCaCu_2O_{6+y}$	1988	105
$TlBa_2Ca_2Cu_3O_{9+y}$	1989	110
$HgBa_2CaCu_2O_{6+y}$	1993	120
$GdFeAsO_{1-y}$	2008	53.5

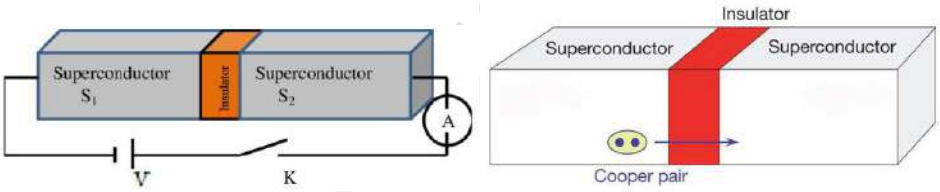
Quantum Tunnelling:

It is a quantum mechanical phenomenon in which an object such as an electron or atom passes through a potential energy barrier. This concept is not possible, according to classical mechanics. Tunneling is an outcome of wave nature of matter and is found in low mass particles like electrons, protons, neutrons etc., Probability of transmission of a wave packet through a barrier decreases exponentially with the barrier height. When a quantum wave reached the barrier, its amplitude will decrease exponentially. Further, this drop in amplitude corresponds to a drop in the probability of finding particle further into the barrier.



In the regions where the potential energy is higher than the waves energy, the amplitude of the wave decays exponentially. If the region is narrow enough, the wave can have a non-zero amplitude on the other side.

Josephson junction:



A Josephson junction is made by sandwiching a thin layer of a non-superconducting material between two layers of superconducting material. The non-superconducting barrier separating the two superconductors must be very thin. If the barrier is an insulator, it must be about 30 angstroms thick or less. If the barrier is another metal, it can be as much as several nanometers thick. In this system, the cooper pairs tunnel through the barrier without resistance. This phenomenon of flow of current between two pieces of superconductor separated by a normal material is called as Josephson effect and the current is called Josephson current. The current flows through the junction even in the absence of external DC voltage. Hence the Josephson current is present in the absence of supply voltage. If the external DC voltage is applied, current oscillates rapidly with a frequency of several GHz, leading to the development of AC voltage.

DC Josephson Effect:

It is the phenomenon of flow of super current through the junction even in the absence of external emf. If the voltage across the junction is measured, it gives zero. Consider a Josephson junction containing two superconducting films separated by thin oxide layer. Here cooper pairs in the superconductor starts tunneling through the oxide layer which are represented by wave function. During this process the oxide layer introduces the phase difference between input and output wave functions. Due to this, super current flow through the junction, even in the absence of external source. The super current through the junction is

$$I_S = I_C Sin\phi_0$$

I_C is critical current at zero voltage, which depends on the thickness of the junction layer and the temperature,

ϕ_0 is Phase difference between the wave functions of cooper pairs

AC Josephson Effect:

When dc voltage is applied across the Josephson junction, it leads to the development of oscillating current. In other words, an alternating emf of high frequency is established across the junction. This effect is called as AC Josephson effect. The oscillating current is because of the fact that, the application of dc voltage across the junction causes the additional phase change for the cooper pairs. The energy difference of cooper pairs on both sides is of the order of 2eV.

Thus the current $I_S = I_C Sin(\phi_0 + \Delta\phi)$

$\Delta\phi$ is the phase difference and I_c is the critical current.

The frequency of the generated AC is

$$f = \frac{2eV}{h}$$

Where 2eV is the energy difference between the cooper pairs on either side of the Junction. Thus, if a voltage of about 1μV is applied then AC frequency of about 484 MHz can be obtained.

$$NOTE: f = \frac{2 \times 1.602 \times 10^{-19} \times 1 \times 10^{-6}}{6.63 \times 10^{-34}}$$

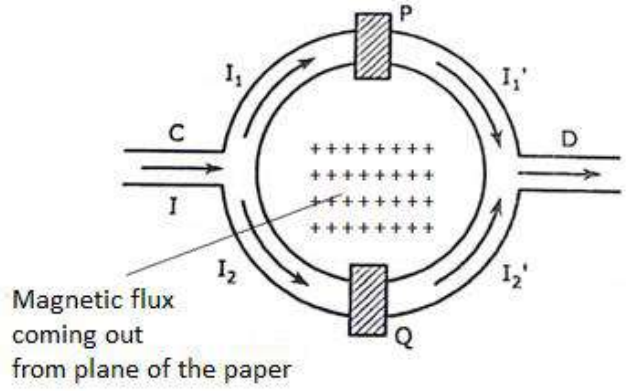
SQUIDS:

SQUID stands for Superconducting Quantum Interference Device. It is an instrument used to measure extremely weak magnetic field of the order of 10^{-13} to 10^{-21} tesla. Hence it is a sensitive magnetometer. Heart of the SQUID is a superconducting ring containing one or more Josephson junctions. Two types of SQUIDS are available. Namely, DC - SQUID and RF - SQUID. They work on the principle of Josephson effect.

DC - SQUID:

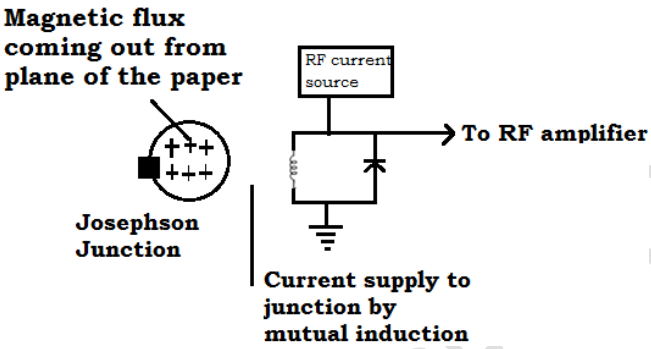
It has two Josephson junctions connected in parallel and works on the interference of current from two junctions. It works on the principle of DC Josephson effect which is the phenomenon of flow of super current through the junction even in the absence of external emf or voltage.

Construction and working:



The cross sectional view of the arrangement is shown. P and Q are two Josephson Junctions arranged in parallel. When current I flows through the point C, it divides into I_1 and I_2. Hence the wave function due to these super currents (cooper pairs) experiences a phase shift at P and Q. In the absence of applied magnetic field, the phase difference between the wave functions is zero. If the magnetic field is applied perpendicular to the current loop, then phase difference between the wave functions will not be zero. This is identified by the sum of the currents I_1 and I_2. The magnitude of phase difference is proportional to applied magnetic field. Hence, even if there is a weak magnetic field in the region will be detected.

RF SQUID:



It works on the principle of AC Josephson effect. When dc voltage is applied across the Josephson junction, it leads to the development of oscillating current. It has single Josephson junction. Magnetic field is applied perpendicular to the plane of the current loop. The flux is coupled into a loop containing a single Josephson junction through an input coil and an RF source. Hence when the RF current changes, there is corresponding change in the flux linked with the coil. This variation is very sensitive and is measured. It is also used in the detection of low magnetic field. It is less sensitive compared to DC - SQUID. Due to its low cost manufacturing, it is commonly used in many applications.

Problem: - A superconducting tin has a critical field of 306 gauss at 0K and 217 gauss at 2K. Find the critical temperature of superconducting tin.

Data: - H_0=306gauss at 0K, H_C=217gaussat 2K=T

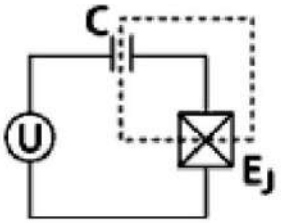
$$H_C = H_0\left(1-\frac{T^2}{T_c^2}\right)$$

$$T_C = \frac{T}{\sqrt{1-\frac{H_c}{H_0}}} = \frac{2}{\sqrt{1-\frac{217}{306}}} = 3.71\,K$$

Applications of Superconductivity is Quantum computing:

SQUIDs find application of magnetometers to measures very small magnetic fields like human brain and heart magnetic fields. But the applications of SQUIDs in quantum computing are as follows.

a) Charge Qubit:



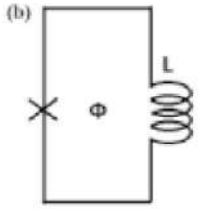
In quantum computing, a charge qubit is also known as Cooper – pair box. It is a qubit whose basis states are charge states. The states represent the presence or absence of excess Cooper pairs in the island (dotted region in the Fig.). In superconducting quantum computing, a charge qubit is formed by a tiny superconducting island coupled by Josephson Junction to a superconducting reservoir.

b) Flux Qubit:

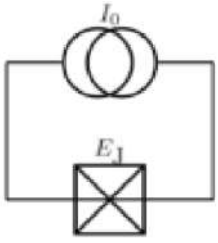
Flux qubits (also known as persistent current qubits) are micrometer sized loops of superconducting metal that is interrupted by a number of Josephson junctions. These devices function as quantum bits. The Josephson junctions are designed so that a persistent current will flow continuously when an external magnetic flux is applied. Only an integer number of flux quanta are allowed to penetrate the superconducting ring.

c) Phase Qubit:

A phase qubit is a current biased Josephson junction operated in zero voltage state with a non-zero current bias.



This employs a single Josephson junction and two levels are defined by quantum oscillations of the phase difference between the electrodes of the junction. DC squid is a type of phase qubit.



Q. No.	Questions
1.	State and explain Meissner Effect.
2.	Define critical field and hence explain its variation with temperature below critical temperature.
3.	Explain Type -1 Type – 2 superconductors with neat figures.
4.	Distinguish between Type -1 Type – 2 superconductors.
5.	Describe high temperature superconductors.
6.	Elucidate the BCS theory of superconductivity.
7.	Explain the phenomenon of quantum tunneling.
8.	Define Josephson junction and hence explain DC and AC Josephson effects.
9.	Define SQUID and describe DC and RF SQUIDs.
10.	Brief out the applications of superconductivity in quantum computing.

Q. No.	Numerical problems
1.	Lead has superconducting transition temperature of 7.26 K. If the initial field at 0 k is 50x103 A/m, calculate the critical field at 6 K.
2.	A superconducting tin has a critical temperature of 3.7 K at zero magnetic field and a critical field of 0.0306 tesla at 0 K. Find the critical field at 2 K.
3.	The superconducting transition temperature of lead is 7.26 K. Calculate the initial field at 0 K, given the critical field at 5 K is 33.644x103 A/m.
4.	Calculate the ratio of critical fields for a superconductor at 7 K and 5K, given the critical temperature is 8 K.
5.	The critical field for niobium is 1x105 A/m at 8 K and 2x105 A/m at 0 K. Calculate the transition temperature of the element.

*** * * * *END* * * * ***

MYSURU ROYAL INSTITUTE OF TECHNOLOGY

DEPARTMENT OF PHYSICS

Applied Physics for CSE Stream : 2022-23

Module – 5 : Module - 5, Applications of Physics in computing

Notes

Syllabus:

Applications of Physics in computing: **8 hrs**

Physics of Animation:

Taxonomy of physics based animation methods, Frames, Frames per Second, Size and Scale, Weight and Strength, Motion and Timing in Animations, Constant Force and Acceleration, The Odd rule, Odd-rule Scenarios, Motion Graphs, Examples of Character Animation: Jumping, Parts of Jump, Jump Magnification, Stop Time, Walking: Strides and Steps, Walk Timing. Numerical Problems

Statistical Physics for Computing:

Descriptive statistics and inferential statistics, Poisson distribution and modeling the probability of proton decay, Normal Distributions (Bell Curves), Monte Carlo Method: Determination of Value of π. Numerical Problems.

Pre requisites: Motion in one dimension, Probability

Self-learning: Frames, Frames per Second

Physics of Animation

Introduction:

Animation is a method of photographing successive drawings or models to create an illusion of movement in a sequence. Since our eyes can retain an image for nearly 0.1 s, when multiple images appear fast, the brain blends them into a single moving image.

In initial days animation was done by drawing or painting pictures on transparent celluloid sheets and then photographed but today most of the animation work is done with computer-generated imagery or CGI.

Animation is inherently fake. It's not real actors on real sets—it is entirely manipulated from a computer. Even though animators sometime break the laws of physics animation requires an understanding of physics and an animator should have a basic understanding of mechanics and bio-mechanics. Principles of physics are universal, they can be applied to cartoon-style drawings as well as CGI and make audiences to escape reality and enter a fantasy world!

The Taxonomy of Physics-Based Animation Methods (classifications)

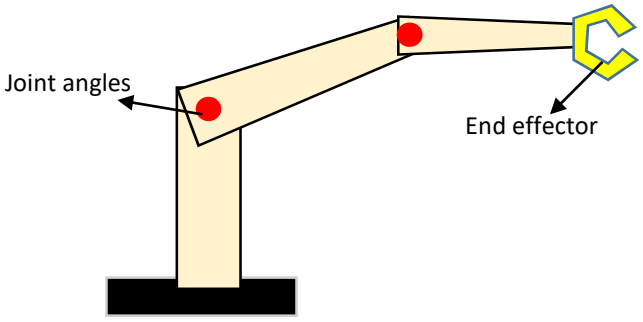
Since animation is an illusion of movement in sequence one need to consider the physics of motion. There are two categories in this. They are

- **Kinematics**

The study of motion of bodies without considering the actual cause for the motion is known as *kinematics*. Here mass and forces are not considered.

- **Dynamics**

The study of motion of bodies by considering the actual cause for the motion is known as *dynamics*. Here mass and forces are taken into consideration



In both of these we have two subgroups. They are

- **Inverse:** It is the study of motion when both the starting and ending points are known. Here one typically knows where to go, but needs to figure out how to do it. It is also known as *backward.* Ex, in robots the motion of its joint angles in robots is inverse or backward

- **Forward:** It is the study of motion when only the starting point is given. Here the goal is to predict the final destination. Ex, end effector movement in the robots is forward motion. It only moves towards the destiny

Frames and Frames per second (FPS)

In animation successive drawings or pictures are made. Each picture is treated as *frame*. To create the appearance of motion from these images, frames shall be displayed.

Definition: The number of frames displayed in one second for smooth movement effect is known as *frame rate*, or *frames per second* (FPS)

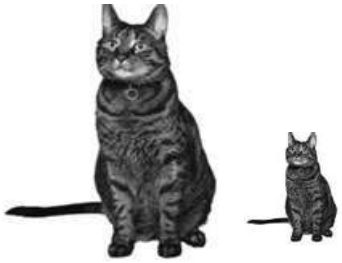
The human brain can process only about 10 to 12 FPS, ie, individuals can distinguish separate still pictures in a series with a frame rate under 12 FPS. 12 frames do produce the motional effect but may look choppy (recall very old black and white movies!!-may be Charlee Chaplin movies). Hence the frame rate has to be increased to produce smooth moving effects. In cinema, a frame rate of 24 frames per second is often used (Different video formats have different FPS). For TV the frame rate is 30 FPS

Scale and Size

Animation of large objects (cricket field, railway stations, dams etc) is not done in their actual size. Their size must be reduced several times. This is nothing but *scaling*. Scaling is not only done for larger dimension objects. Even smaller dimensional objects (insects, body organs etc) also need scaling.

Definition: The ratio of animation size to the actual size of the object is known as **scale** or **scale of animation**. It is also defined as the ratio of pixels of the animation to the physical units of length.

Scaling is not only done for altering the sizes of the characters but also done in other aspects such as movement, energy etc. This is essential to distinguish between other parameters such as weight and strength, younger and older etc.



Improper scaling

Proper scaling

For ex: If only the size of the cat and its kitten is scaled down then the kitten will look like small adult cat only. Actually kitten will move slower than cat and has lesser strength. Hence the scaling is needed to these factors also for better effects

Proportion and Scale (weight and strength)

When object is scaled its volume and area does not change in equal proportionate. Change in the volume is more compare to area. Volume increases by *cube* times and area increases by *square* times. Hence proportionate scaling is very important. Consider the following example.

The **weight** of a man depends on his body volume and the muscle **strength** depends on cross sectional area. If we want to double the muscle strength the width must be increased $\sqrt{2}$ times. But the weight is not scaled proportionately. Hence care has to be taken while creating larger or smaller objects by means of scaling.

Motion and Timing in Animations

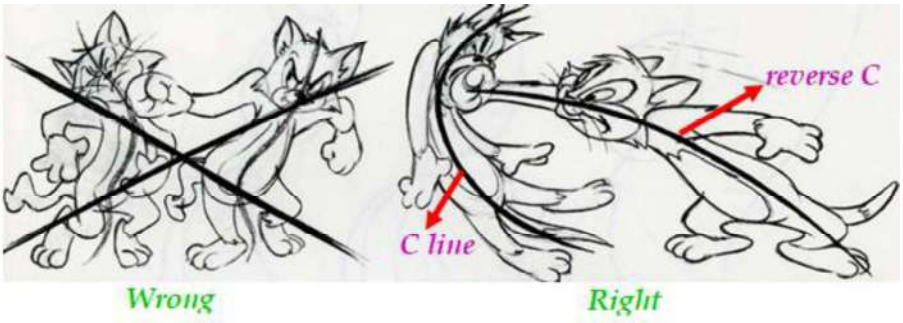
Motion is an essential part of animation. Several types of motion need to be considered while animating a scene. Common types of motion are

- Linear motion
- Parabolic motion
- Circular motion
- Wave motion or oscillatory motion

Line of action and path of action

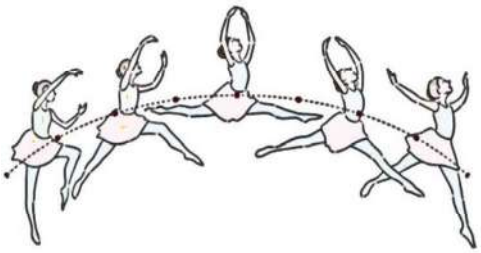
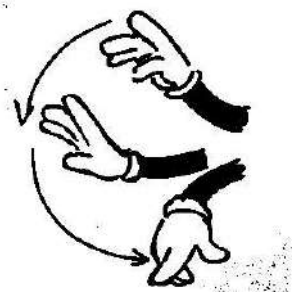
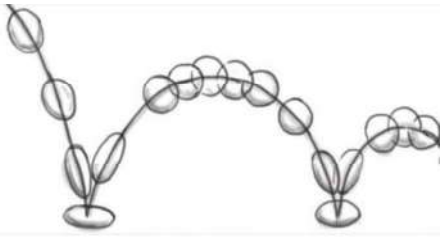
All these types are familiar to us. In connection with animation of motion, we define two parameters namely *line of action* and *path of action*.

Line of action is an imaginary line that describes the direction and motion of a character's body. A good line of action improves character's poses, makes the character look more dynamic, energetic and alive. There are 3 types of lines of action, the C, reverse C and S curves. Look at the following fig. It is clear that the second pose gives better impression than first one



The path along which the object or character moves is known as ***path of action*** or ***arc.*** For ex, path of bouncing ball, trajectory of projectile, jumping of ballet dancer, path of moving arm etc.

Arcs are used extensively in animation because they create motion that is more expressive and less stiff than action along a straight path.

Timing

Timing refers to the time it takes for an action to complete from the starting point to the end. Timing is the amount of frames it takes for an action to take place.

Timing can have a huge effect on how we perceive a character. If a character moves with less frames rate it will appear to be heavy and likely very big. If a character has high frames rate it will appear to be light and probably small. It is explained in the following example

Consider three blocks of same size. 1st block is given 6 frames rate, 2nd block is given 12 frames rate and 3rd one is given 24 frames rate. As a result 1st block falls fast giving the illusion of being heavy

Timing of action consists of placing objects or characters in particular locations at specific frames to give the illusion of motion. If images are placed at same location in all the frames it won't give motion effect (fig-A). Hence we work with very small intervals of time (fig-B).

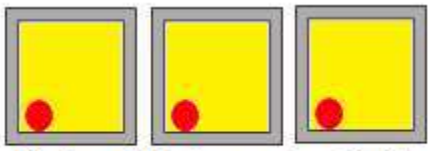


Fig-A- red dot at same place in all the frames

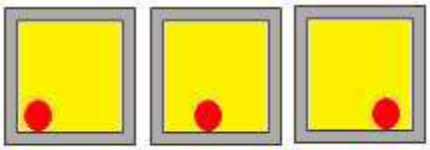
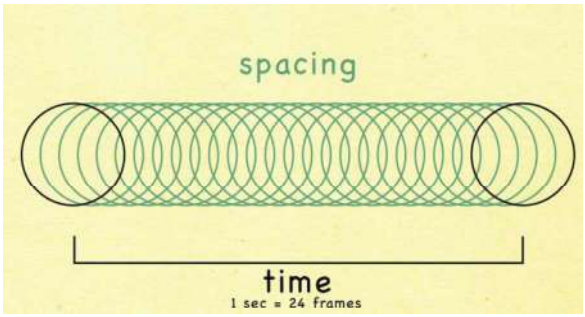


Fig-B-red dot at different places

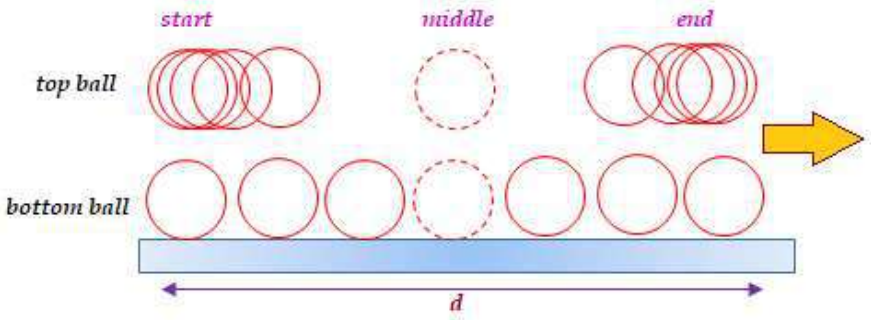
Spacing

Spacing is the distance an object moves within a defined time for that action (it is also the distance an object moves for every frame of that action)



The main difference between timing and spacing is that timing refers to everything that happens over time whereas spacing refers to how much fast something moves. It is illustrated as follows

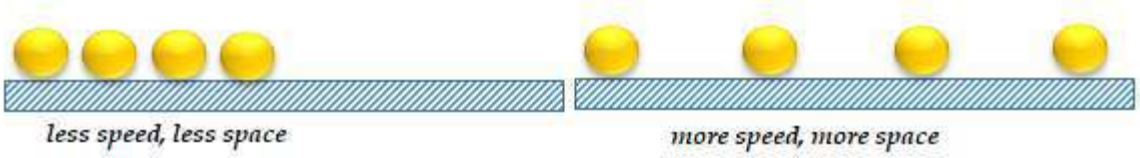
Consider two balls moving same horizontal distance in same time. Assume the top ball will be moving slowly to begin with, then speed up and then slow down again at the end. The bottom ball will just begin and keep a constant speed until the end. Both will be in the same position at the beginning, middle and end of the animation. To achieve this we are going to change the spacing. If we have more drawings near the starting pose, one or two in the middle, and more drawings near the next pose lesser drawings make the action faster and more drawings make the action slower.



Linear motion, Uniform motion and timing

A body moving in a straight line is said to be in *linear motion*. It always move along same direction. Ex: a ball rolling on the inclined plane, a ball moving on a horizontal plane, a stone falling under gravity etc

In linear motion if the body travels without acceleration (or deceleration) then it is known as *uniform motion.* Here the net force acting on the object is zero and distance travelled between consecutive frames is same. Hence timing is very easy. If the speed is less the spacing between the frames is less. If the speed is more the spacing is also more



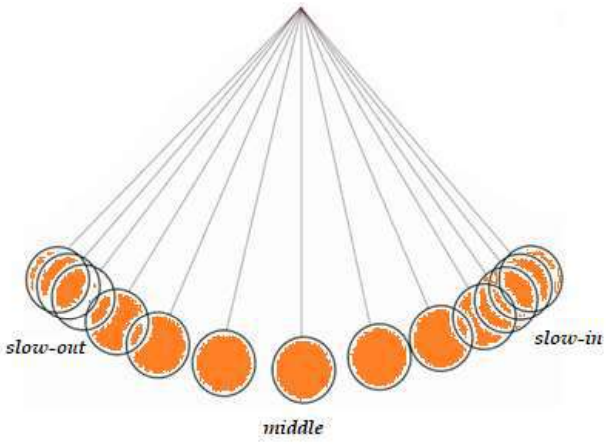
When the net force acting on the object is not equal to zero then according to Newton's second law, there will be acceleration (even deceleration also). As a result the object will get faster and faster until the force stops acting on it. Now comes one of the Principles of Animation. It is called "slow-in and slow-out".

Slow-in and slow-out

All types of movements start with acceleration and end with deceleration. Even in animation also this principle is followed. An object or a character in animated video starts its movements more slowly, then picks up speed and finishes with deceleration. As a result the beginning and end of the movement are softened. This is called 'slow-in' (or ease in) &'slow-out'(or ease out)

.
Slow-in (or ease in) is process in which the body is preparing for stopping and **Slow-out** (or ease out) is the process in which the body is speeding up from a still position.

Slow-in and slow-out is achieved by adjusting the spacing (as explained in the above section). To create a slow-out, we need to place the frames close together and then move them apart gradually. To create the slow-in we will need to gradually place the frames closer together at the end of the animation.



Consider the animation of oscillating pendulum. There are more circles at slow-in and slow-out positions and less at middle positions.

Constant force, acceleration and timing

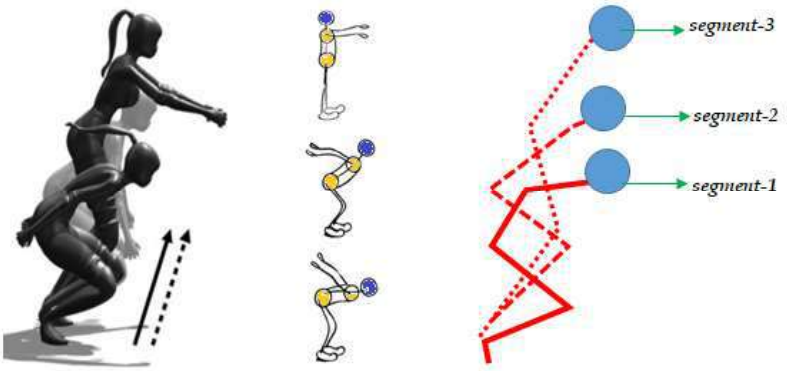
If the applied force is not varying with time then it is called constant force. Some of the examples for constant force are *gravitational force, frictional force, force on a charge in uniform electric field, force on a charge in uniform magnetic field etc.*

If the force is varying with respect to time then it is variable force. Ex, *force on a charge in variable magnetic and electric field, spring force etc.* Timing of the action is very difficult for motions due to variable forces.

Timing for variable force

Normally all the forces exerted by us in everyday life are variable. Force applied during walking, running, jumping, skating, dancing etc is a variable one. If you break the motion into smaller segments then force in each segment is treated as constant and timing becomes easier.

In the below fig, jumping (by girl) is divided into 3 segments. Force is nearly constant over each segment. The timing is very short for each of these segments



Timing for constant force and acceleration

If the applied force is constant then the body will move with constant acceleration. The acceleration depends on the direction of the applied force. There are 3 possibilities

i. If a constant force is applied on a body at rest then it accelerates
ii. If a constant force is applied in the direction of motion then the body is accelerated
iii. If a constant force is applied in the opposite direction to motion then it is decelerated.

For such type of motions timing is done with **odd rule**

Odd rule

The distance travelled by the object between two successive frames is calculated by odd rule. According to this rule

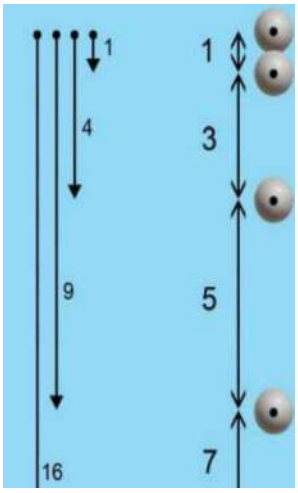
"The distance travelled between two successive frames is obtained by multiplying the base distance by odd numbers 1,3,5,7 etc during accelerated motion and by multiplying the base distance by odd numbers 7,5,3,1 etc during decelerated motion"

Base distance: Base distance is the smallest distance between the two frames. For a slow-out, this is the distance between the first two frames; for a slow-in, it's the distance between the last two frames. It is given by

$$\text{Base distance} = \frac{\text{total distance}}{(\text{last frame number} - 1)^2}$$

Explanation:

This rule is applicable to both vertical and horizontal motion. It is explained for a body falling under gravity. Consider a ball falling under gravity with zero initial velocity. Distance travelled



$$\text{after 1 sec} \qquad S_1 = \frac{1}{2}gt^2 = \frac{1}{2}g = 1 \times \left(\frac{1}{2}g\right)$$

$$\text{after 2 sec} \qquad S_2 = \frac{1}{2}gt^2 = 4 \times \left(\frac{1}{2}g\right)$$

$$\text{after 3 sec} \qquad S_3 = \frac{1}{2}gt^2 = 9 \times \left(\frac{1}{2}g\right)$$

$$\text{after 4 sec} \qquad S_4 = \frac{1}{2}gt^2 = 16 \times \left(\frac{1}{2}g\right)$$

and so on

Here the difference between any two successive distances is an odd number. For ex, 4-1 = 3; 9-4 =5; 16-9 = 7 and so on. Hence the rule is named as odd rule

Odd rule multipliers

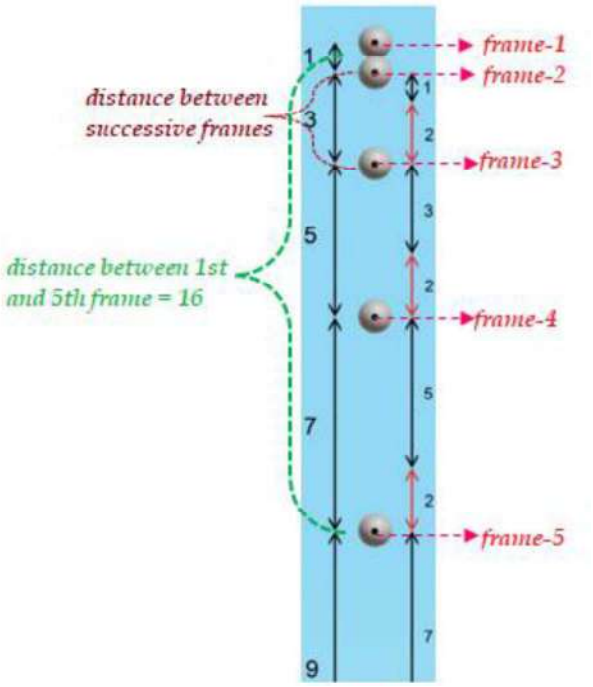
The process of calculating the distance from the first frame to the current frame and use these distances to place the object on specific frames is known as **odd rule multiplier**. For slow-out, it is as follows

i. Find the distance between first two frames. It is also known as "base distance"
ii. Multiply it by 1 to get the distance between frame 1 and 2. The total distance travelled is 0+1=1
iii. Multiply by 3 to get the distance between frame 2 and 3. The total distance travelled is 0+1+3=4
iv. Multiply by 5 to get the distance between frame 3 and 4. The total distance travelled is 0+1+3+5=9

And so on…It is summarized as follows

Frame No. (A)	Distance between successive frames (B)	Distance between 1st to present frames (C)
1	--	0
2	1	$0+1=1=1^2$
3	3	$0+1+3=4=2^2$
4	5	$0+1+3+5=9=3^2$
5	7	$0+1+3+5+7=16=4^2$
6	9	$0+1+3+5+7+9=25=5^2$

The same is represented in the following fig. From fig, it may be observed that after the first increment, all the others are longer by the same amount (amount of 2- Red arrow mark).
Ie, 3 – 1 =2; 5 – 3 = 2; 7 – 5 = 2 and so on



It is possible to write the relation between A, B and C as follows

$$B = 2(A-1) - 1$$

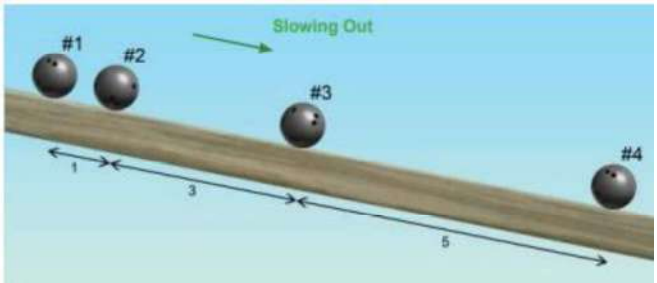
$$C = (A-1)^2$$

Ex:

If A = 5 then B = 2(5-1)-1 = 7

$$C = (5-1)^2 = 16$$

This is true even for a horizontal motion also



The whole process is applicable for 'slow-in' process also (decelerated motion). Only difference is that we have to multiply in the reverse manner ie, 7, 5, 3, 1.

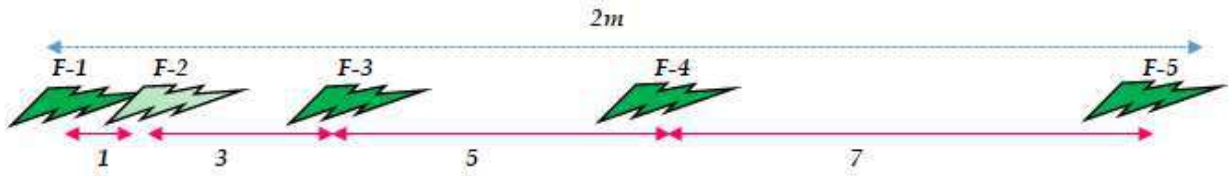
Odd Rule Scenarios

Here we summarize the odd rule for 4 different cases.

1. *Slow-out process*

First of all, find the distance between first two successive frames. It is also known as *base distance*. Then multiply it by odd numbers 1,3,5,7 etc to get the distances between consecutive frames. Finally use squares to multiply the base distance to get the total distance traveled on each frame (See the above table)
If the total distance is known then the base distance is calculated using the formula

$$\text{Base distance} = \frac{\text{total distance}}{(\text{last frame number} - 1)^2}$$



For ex: In the above fig, the total distance is 2m and number of frames is 5 then

$$\text{Base distance} = \frac{2}{(5-1)^2} = 0.125m$$

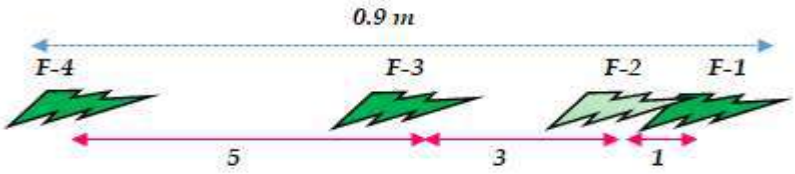
Distance between frame 1 and 2 = 1 × 0.125m = 0.125 m
Distance between frame 2 and 3 = 3 × 0.125m = 0.375 m
Distance between frame 3 and 4 = 5 × 0.125m = 0.625 m
Distance between frame 4 and 5 = 7 × 0.125m = 0.875 m

∴ Total distance = 0.125+0.375+0.625+0.875 = 2m

2. *Slow-in process*

First, find the distance between the last two frames. Then multiply it by odd numbers in reverse order, ie 7,5,3,1
If the total distance is known then the base distance is calculated using the formula

$$\text{Base distance} = \frac{\text{total distance}}{(\text{last frame number} - 1)^2}$$



In the above fig, the total distance is 0.9 m and number of frames is 4 then

$$\text{Base distance} = \frac{0.9}{(4-1)^2} = 0.1m$$

Distance between frame 4 and 3 = 5 × 0.1m = 0.5 m
Distance between frame 3 and 2 = 3 × 0.1m = 0.3 m
Distance between frame 2 and 1 = 1 × 0.1m = 0.1 m

∴ Total distance = 0.5 + 0.3 + 0.1 = 0.9 m

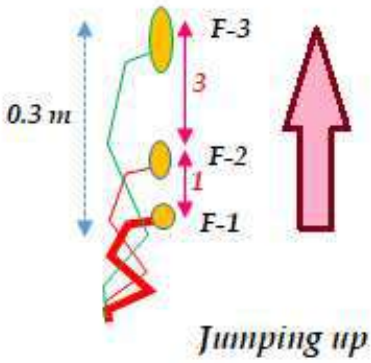
3. *Jumping up process*

Consider a girl jumping up. The whole process is divided into 3 small intervals (or 3 frames). Let the total height is 0.3 m, then

$$\text{Base distance} = \frac{0.3}{(3-1)^2} = 0.075m$$

Distance between frame1 and 2 = 1 × 0.075m = 0.075 m
Distance between frame2 and 3 = 3 × 0.075m = 0.0225 m

∴ Total distance = 0.075 + 0.225 = 0.3 m



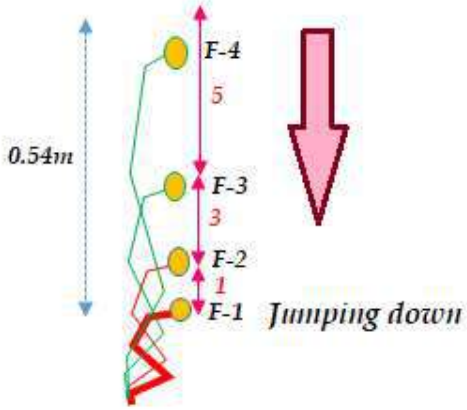
4. *Jumping down process*

Consider a girl jumping down through a distance 0.54 m. The whole process is divided into 4 frames, then

$$\text{Base distance} = \frac{0.54}{(4-1)^2} = 0.06m$$

Distance between frame 4 and 3 = 5 × 0.06m = 0.3 m
Distance between frame 3 and 2 = 3 × 0.06m = 0.18 m
Distance between frame 2 and 1 = 1 × 0.06m = 0.06 m

∴ Total distance = 0.3 + 0.18 + 0.06 = 0.54 m



Motion Graphs

A graph of distance versus time is known as **motion graph**. It is also called as position-time graph or *x-t* graph. It is very important for animation because the nature of the motion is ascertained from the graph. Some of the motion graphs are as follows

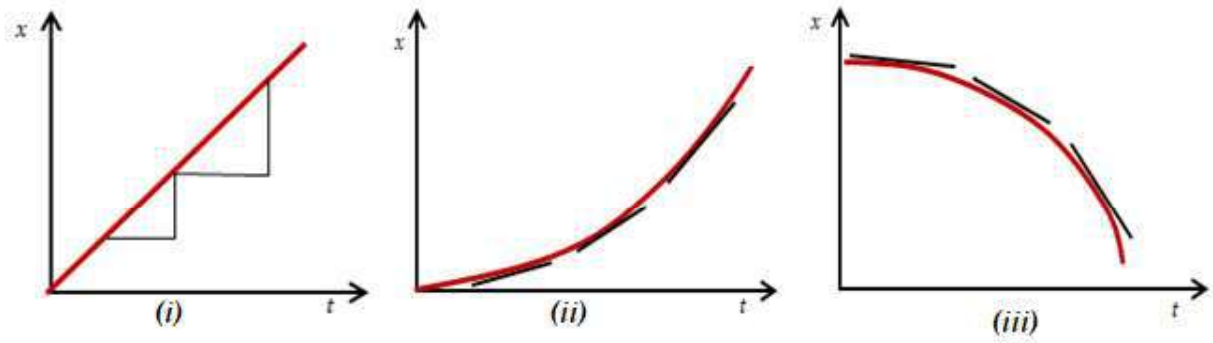


Fig. (i) represents constant velocity and constant slope. There is no acceleration
Fig. (ii) represents positively increasing slope and velocity. Hence acceleration is positive. More the acceleration more will be the curvature
Fig.(iii) represents negatively increasing slope. It represents decreasing velocity. Hence it is deceleration

Examples of Character Animation

Motion of an object is an integral part of animation. By observing and studying the real life movements one can make the animation more attractive and lively. Jumping and walking are two real life actions where the entire body is in motion. Now we shall take up their character animation by applying the knowledge of physics in animation

1. Jumping

Following are the different stages of jumping process.

- ***Crouch***—It is the bending pose taken as preparation for jumping.

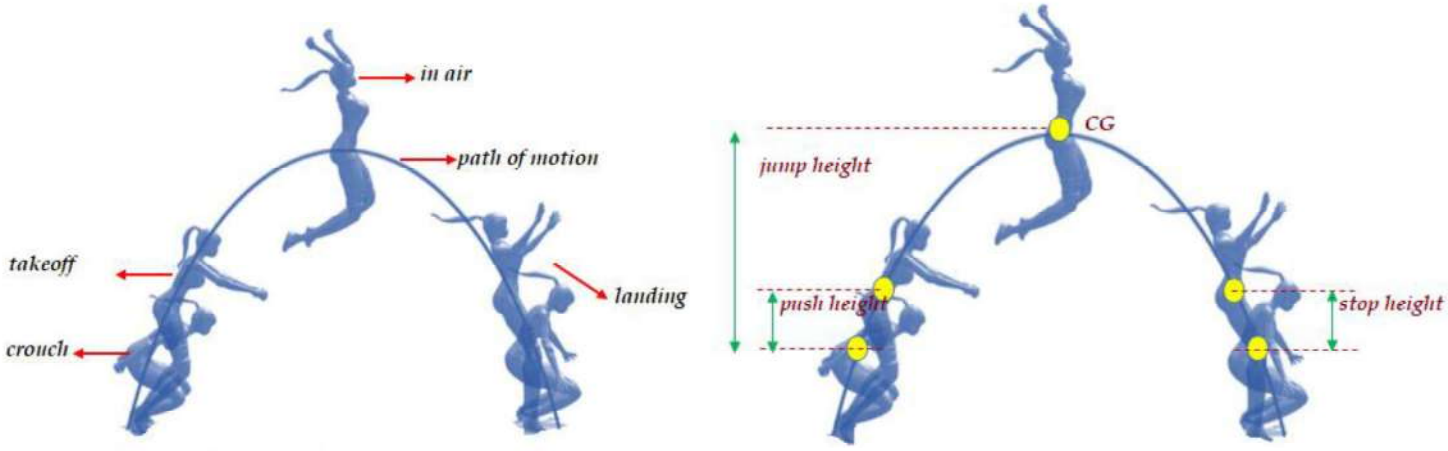
- ***Takeoff*** —Character pushes up fast and straightens legs with feet still on the ground. The distance from the character's center of gravity (CG) in the crouch to the CG in takeoff position is called the *push height*. The amount of time (or number of frames) needed for the push is called the *push time*.

- ***In the air***—Both the feet are off the ground, and the character's center of gravity (CG) moves in a parabolic arc. First it reaches a maximum height (apex), and then falls back to the ground. The maximum height or *jump height*, is measured from the CG at takeoff to the CG at the apex of the jump. The amount of time the character is in the air from takeoff

to apex is called the *jump time*. If the takeoff pose and the landing pose are similar, then jump time is same for going up and coming down (it is exactly similar to time of ascent and time of descent in projectile motion).

- **Landing**—Character touches the ground and bends knees to return to a crouch. The distance from the character's CG when her feet hit to the ground to the point where the character stops crouching is called the *stop height*. The stop height is not always exactly the same as the push height

- **Path of action**—It is the trajectory along which the character jumps



Calculations

i. Calculation of jump timing

Let the jump height be 1.2 m and acceleration due to gravity is 9.8 ms^{-2}. The jump time is

$$t^2 = \frac{2h}{g} = \frac{2 \times 1.2}{9.8} = 0.25$$

$$\therefore t = 0.5\ s$$

This is the time taken to go from crouch position to apex position. As we know that the frame rate is 24 (FPS), the total number of frames for this action should be $24 \times 0.5 = 12$ frames

ii. Jump magnification (JM)

The JM is the ratio of the jump height to the push height.

$$JM = \frac{\text{jump height}}{\text{push height}} - - - (1)$$

It is also defined as the ratio of jump time and push time. It is used to calculate the push timing and stop timing.

$$JM = \frac{\text{jump time}}{\text{push time}} - - - (2)$$

In terms of acceleration it is given by

$$JM = \frac{\text{push acceleration}}{g} - - - (3)$$

[To prove Eq (3): (Optional – not in syllabus)

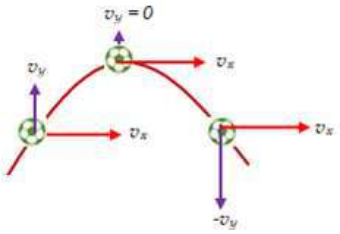
Jumping motion is equivalent to projectile motion. While going from crouch to take off position the girl will have some acceleration called *push acceleration*. After wards her body moves upward and experiences acceleration due to gravity. During downward motion she experiences gravity only.

Jump velocity = push acceleration × push time (because $v = at$)

Landing velocity = landing acceleration × jump time

But the jump velocity shall be equal to landing velocity. Hence

push acceleration × push time = landing acceleration × jump time



Since landing acceleration is same as g we can write

push acceleration × push time = g × jump time

$$\frac{\text{jump time}}{\text{push time}} = \frac{\text{push acceleration}}{g}$$

$$JM = \frac{\text{push acceleration}}{g}$$

Hence the proof]

Ex, if there are 15 frames and JM is 3 then

According to Eq (2)

$$JM = \frac{\text{jump time}}{\text{push time}} = \frac{15}{\text{push time}}$$

$$\text{push time} = \frac{15}{3} = 5 \text{ frames}$$

iii. Jump height

The distance between CG at crouch and take off position is called push height. Assume that it is 0.4 and JM is 3 then from eq (1) we get

$$JM = \frac{\text{jump height}}{\text{push height}} = \frac{\text{jump height}}{0.4}$$

$$\text{jump height} = \text{JM} \times 0.4 = 3 \times 0.4 = 1.2 \text{ m}$$

iv. Stop time

Since timing of the push and stop are same we can write

$$\frac{\text{push height}}{\text{push time}} = \frac{\text{stop height}}{\text{stop time}}$$

$$\text{stop time} = \text{stop height} \times \frac{\text{push time}}{\text{push height}}$$

It may be noted here that the stop height is little bit longer than push height (because the girl may finally lay down on the ground after landing). For ex, if push height is 0.4, stop height is 0.5, push time 5 frames then

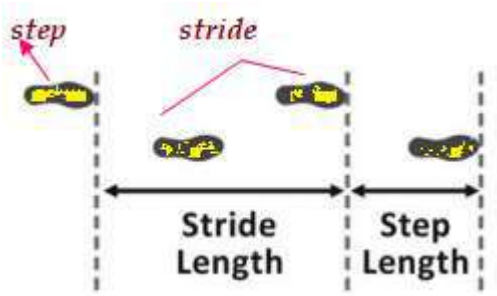
$$\text{stop time} = 0.5 \times \frac{5}{0.4} \approx 6\ frames$$

2. Walking

Walking is another event where a lot of physics is involved and the animation of walking needs the complete knowledge of mechanics. Walking is nothing but a step-by-step movement. Each step is divided into 4 poses namely **passing, step, contact,** and **lift**. Also we need the concept of Strides and steps as well as walking time

Strides and Steps

A step means one step with one foot. A stride means two steps, one with each foot. *Step length* is the distance between two successive steps and *Stride length* is the distance between two parts of the same step. Step and stride length indicate lengthwise spacing for the feet during a walk.



Walk timing

While walking each foot is in contact with ground for 60% of time (ie, single support) and both the feet will be in contact for 20 % of time (double support). In this connection we define a quantity called *Gait*. It is the timing of the motion for each foot, including how long each foot is on the ground or in the air. For normal walk it is about ½ second

To walk faster, obviously one has to increase stride length and decrease the time of double support. A fast walk has a stride rate of about 4 feet/sec. If it becomes 6-7 then walking turns into running.

Statistical Physics

Introduction

Statistics is the area of applied mathematics that deals with the collection, organization, analysis, interpretation, and presentation of data.

For ex, in the study of *tiger population* in a particular forest region, one has to use the statistical methods such as **collecting** all possible data (like number of tigers, deer, monkeys, area of the grass field etc) over definite period of time, **arrange** and **compare** it with the previous year's data, identify the differences and **analyze**, find out the reasons for the population variation, **interpret** the result and **present** it in a convincing manner

Statistical Physics

Statistics is equally important in Physics and it is called as **statistical physics**. It is basically evolved from statistical mechanics

Statistical Physics is mainly intended to explain the properties of matter in total, in terms of physical laws governing atomic motion. It also explain numerous physical phenomena, such as superconductivity, magnetism, changes in states of matter, energy state distribution, electron distribution in metals, radioactivity, noise in electronic devices etc,

[Kinetic theory of gases is one of the best example in which macroscopic parameters such as kinetic energy, pressure and temperature are explained in terms of microscopic quantities-number of molecules and their rms velocity]

Descriptive statistics and inferential statistics

Statistical techniques can be categorized as descriptive statistics and inferential statistics

The process of organizing and describing the known data using charts, bar graphs, etc., is known as **descriptive statistics.** It mainly focuses on describing the visible characteristics of a database. It enables researchers to present data in a more meaningful way such that easy interpretations can be made.

The process of making predictions or inferences (conclusions) and generalizations about the data is known as **inferential statistics.** When the data is very large it becomes difficult to use it. In such cases, certain samples are taken as representative of the entire population. Inferential statistics draws conclusions using these samples.

Both descriptive and inferential statistics need to be used hand in hand so as to analyze the data in the best possible way.

Ex:
Assume that you have the marks scored by 100 students (of a specific school where the total strength is 5000).
By using descriptive statistics, the spread of the marks can be done and can obtain a clear idea regarding the performance of each student.
Inferential statistics will now make use of this sample data of 100 students and make generalizations about the population (ie all 5000 students).

Poisson's distribution

Introduction
Look at these examples

i. Count the number of accidents occurring in a metropolitan city in one year
ii. Count the number of deaths occurred in frequently flooded area
iii. Count the number of people visit a shopping mall a day
iv. Count the number of radioactive nuclei disintegrate in a sample
v. Count the number of collisions between electrons and lattice points in a sample of given volume
vi. Count the number of website visitors per month

In all these examples we are **counting** some of the **events** (green color) occurring within a given interval of time or space or volume. Events are occurring randomly and independently. By knowing 'when one event happens' gives absolutely no information about 'when another event will occur'. For ex, if one U-235 nuclei disintegrates now, it is impossible to predict when the other one will disintegrate (same thing is true in all the above examples). In such cases the probability of a given number of events occurring in a fixed interval of time or space is explained using **Poisson's distribution function**

Definition and explanation

Poisson's distribution is a discrete distribution and describes probabilities for counts of events occurring in a specific interval of time or space. Counts may be random but over certain interval of time it will be a 'definite average'. The discrete outcome is a number represented by k. It is always non negative whole number.

The Poisson distribution mass function is given by

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where

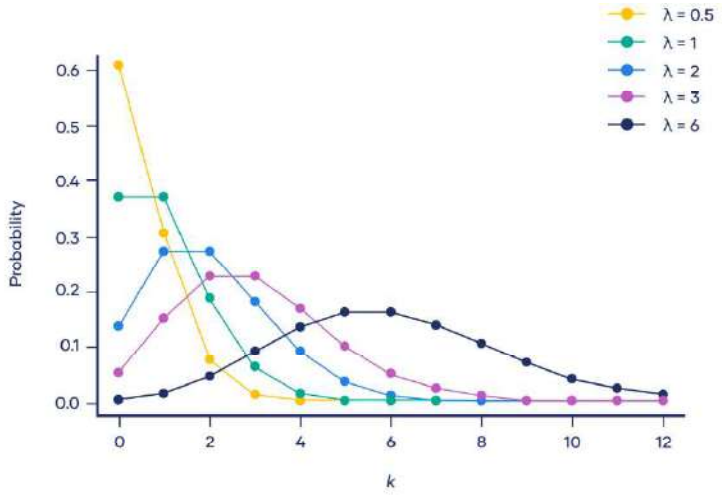
X is a random variable following a Poisson distribution
k is the number of times an event occurs in given interval of time or space
P is the probability that an event will occur *k* times
λ is the average number of times an event occurs

Ex:
One of the first applications of the Poisson distribution was by statistician Ladislaus Bortkiewicz. In the 1870s he investigated accidental deaths by **horse kick** of soldiers in the Prussian army. He analyzed 20 years of data for 10 army corps (batch), (equivalent to 200 years of observations of one corps)
He found that a mean of **0.61 soldiers per corps** died from horse kicks **each year**. However, most years, no soldiers died from horse kicks. On the other end of the spectrum, **one tragic year** there were **four soldiers** in the same corps who died from horse kicks.
Using modern terminology:

- A death by horse kick is an "event."
- The time interval is one year.
- The mean number of events per time interval, λ, is 0.61.
- The number of deaths by horse kick in a specific year is *k (=4)*.

The Poisson distribution may be useful to model events such as the radioactive decays, number of laser photons hitting a detector in a particular time interval, thermionic emission, scattering of particles etc

A graph of P v/s *k* is known as Poisson distribution curve. It is as follows. When λ is low, the distribution is much longer on the right side of its peak than its left. As λ increases, the distribution looks more and more similar to a normal distribution



Ex.1. On a particular river, overflow floods occur once every 100 years on average. Calculate the probability of k = 0, 1, 2, 3, 4, 5, or 6 overflow floods in a 100 year interval, assuming the Poisson model is appropriate.
Ans:

λ = 1 (ie the average flood over 100 years occurs only once)

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Let $k = 0$

$$P(X = 0) = \frac{1^0 e^{-1}}{0!} = 0.368$$

Let $k = 1$

$$P(X = 1) = \frac{1^1 e^{-1}}{1!} = 0.368$$

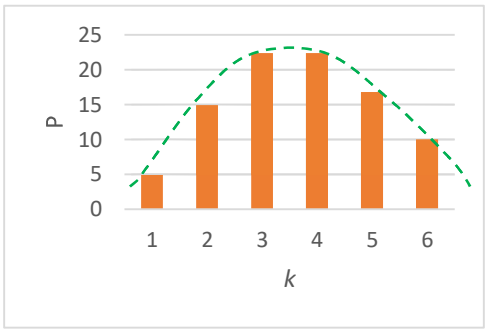
Let $k = 2$

$$P(X = 2) = \frac{1^2 e^{-1}}{2!} = 0.184$$

Ex.2. Careful measurements have established that a sample of radioactive thorium emits alpha particles at a rate of 1.5 per min. what is the expected average result in 2 min interval? What is the probability of getting this number? What is the probability for observing *k* particles for *k* = 0, 1,2,3,4 & 5? (Here *k* is the number of counts or number of particles decay in any 2 min interval)

Here $\lambda = 1.5 \times 2 = 3$

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$



Let $k = 0$ *(ie, no decay in 2 min interval)* and $\lambda = 3$

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = 0.049 = 4.9\%$$

Let $k = 1$ *(ie, 1 decay in 2 min interval)* and $\lambda = 3$

$$P(X = 1) = \frac{3^1 e^{-3}}{1!} = 0.149 = 14.9\%$$

Let $k = 2$ *(ie, 2 decay in 2 min interval)* and $\lambda = 3$

$$P(X = 2) = \frac{3^2 e^{-3}}{2!} = 0.224 = 22.4\%$$

Let $k = 3$ *(ie, 3 decay in 2 min interval)* and $\lambda = 3$

$$P(X = 3) = \frac{3^3 e^{-3}}{3!} = 0.224 = 22.4\%$$

Let $k = 4$ *(ie, 4 decay in 2 min interval)* and $\lambda = 3$

$$P(X = 4) = \frac{3^4 e^{-3}}{4!} = 0.168 = 16.8\%$$

Let $k = 5$ *(ie,5 decay in 2 min interval)* and $\lambda = 3$

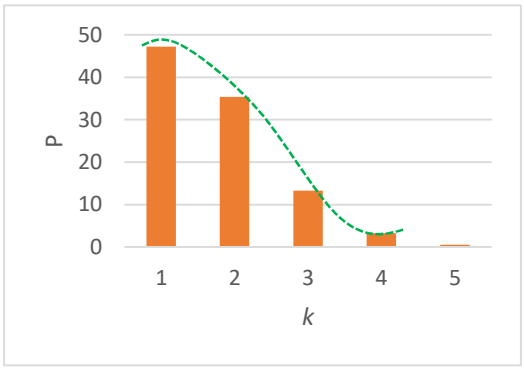
$$P(X = 5) = \frac{3^5 e^{-3}}{5!} = 0.100 = 10\%$$

These results are plotted as shown

Ex:3 On average each of the 18 hens lays 1 egg per day. If you check the hens once an hour and remove any eggs that have been laid what is the average number λ of eggs you find on hourly visit?. Use Poisson distribution to calculate probability that you find *k* eggs for *k* = 0, 1 , 2, 3 and 4

Here 18 hens lay one egg each in 24 hours. Hence the average count per hour is 18/24 = 0.75
ie λ = 0.75

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$



Let *k = 0 (ie, no egg in 1hr interval)* and λ = 0.75

$$P(X = 0) = \frac{0.75^0 e^{-0.75}}{0!} = 0.472 = 47.2\%$$

Let *k =1 (ie, 1 egg in 1hr interval)* and λ = 0.75

$$P(X = 1) = \frac{0.75^1 e^{-0.75}}{1!} = 0.354 = 35.4\%$$

Let *k =2 (ie, 2 egg in 1hr interval)* and λ = 0.75

$$P(X = 2) = \frac{0.75^2 e^{-0.75}}{2!} = 0.133 = 13.3\%$$

Let *k =3 (ie, 3 egg in 1hr interval)* and λ = 0.75

$$P(X = 3) = \frac{0.75^3 e^{-0.75}}{3!} = 0.033 = 3.3\%$$

Let *k =4 (ie, 4 egg in 1hr interval)* and λ = 0.75

$$P(X = 4) = \frac{0.75^4 e^{-0.75}}{4!} = 0.006 = 0.6\%$$

Modeling the Probability for Proton Decay

Background:

In Physics there are some theoretical conservation laws which forbid certain specific reactions or decays to occur in nature. Many Experiments are intended to check the validity of these laws by searching for the presence (or absence) of such reactions or decays. If one or more events are observed in time T interval then the theoretical law is disproven. If no events are observed, the converse cannot be said to be true. Instead, a limit on the life-time of the reaction or decay is set. (This is a sort of Null Experiment aimed at setting confidence limits when no counts are observed). One such example is life time of proton or proton decay

Proton decay

Proton decay is a theoretical form of particle decay. Here proton disintegrates into two particles namely pion and a positron with mean life of 10^{33} years which is very much greater than the age of the universe!. (Age of the universe is about 1×10^{10} years). If you have only one proton then you will have to wait for 10^{33} years to see the decay. But we know the decay is a statistical process. If there are huge number of protons then some of them may decay (if at all the decay is possible!)

Actually conservation laws forbid the proton decay. Despite many serious experiments and significant effort across the globe, even a single proton decay has never been observed so far. One such experiment is explained here

In Super-K detector (JAPAN) 50,000 tons of ultra-pure water containing 7×10^{33} protons is used. The 40% of the experimental area is covered by detector tubes to count the 'event'-decay

Consider the decay law

$$N = N_0 e^{-\lambda t}$$

Where

N_0 = initial number of protons at $t = 0 = 7\times10^{33}$
N = number of protons remain un decayed after t years
$(N_0 - N)$ = number of protons decayed in t years
λ = decay constant = $1/\tau$ (not to be confused with λ in Poisson's formula) = $1/10^{33} = 10^{-33}$

Since λ is very small we can approximate $e^{-\lambda t} = 1 - \lambda t$

$$N = N_0(1 - \lambda t) = N_0 - N_0\lambda t$$

$$(N_0 - N) = N_0\lambda t$$

Put $N_0 = 7\times10^{33}$, $t = 1$ year and $\lambda = 10^{-33}$ we get

$$(N_0 - N) = 7 \times 10^{33} \times 10^{-33} = 7$$

So, 7 decays (events) shall occur in every one year. Since there are 40% detectors, the average detected count must be 40% of 7 ie 2.8 or nearly 3. Hence the mean number of events = $\lambda = 3$. Actually Super-K has started measurement since 1996 but not observed any evidence of proton decay yet. Hence $k = 0$.
Substitute $\lambda = 3$ and $k = 0$ in Poisson's formula we get

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = 0.05$$

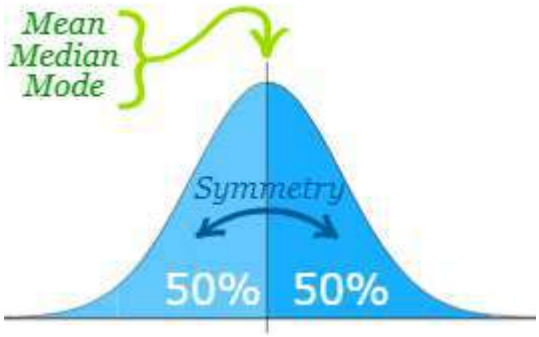
From this low probability for null result (i.e. no decay) Super-K scientists have come to a conclusion that the mean life of proton should be at least 2×10^{34} years! Does it mean that protons live longer than estimated time? or, doesn't decay? Nobody knows the answer! Many countries have taken up this as a challenge and conducting deep underground neutrino experiment (Including India- long back such kind of experiment was initiated near KGF)

Normal distribution (Gaussian distribution) and Bell curve

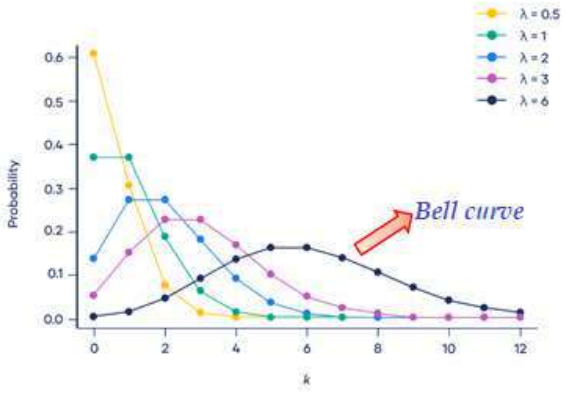
The Normal distribution is one in which the data tends to be distributed equally around a central value with no bias left or right. It is the most universal distribution and plays a central role in all of statistics. The distribution curve assumes bell shape and hence the curve is called "Bell curve". The peak of the curve represents the mean or median

Examples for normal distribution are

- heights of people
- size of things produced by machines
- errors in measurements
- blood pressure
- marks on a test



The Poisson distribution is applied to experiments where the data is strictly bounded on one side; the curve of the graph is highly asymmetric. It is approximated to normal distribution when P, the probability, and λ, the mean number of events are large. The Poisson distribution becomes more symmetric and assumes bell curve shape (as shown in fig)



Monte Carlo Simulation

Introduction

In certain experiments the possible outcome is random and cannot be easily predicted. The outcome takes different values in each experimental trial. Hence the outcomes are called ***random***

variables. Random variables are always real numbers. The likelihood that any of these possible values would occur is known as *probability distribution*.

The concept of *random variable* and *probability distribution* are better understood through these examples

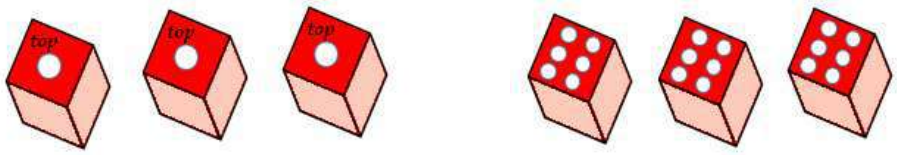
1. Consider the example of **tossing of two coins**. The probability of getting HEADS (or even TAILS) is a random event and it is not likely to happen equally in all the trials. If we take Y as the random variable (ie, getting H) it is either 0 (ie, no head), 1 (ie, one head) or 2 (ie two heads). Since the coins can land in 4 possible ways namely TT, TH, HT and HH then probability distribution of

   i. P(Y = 0) = 1/4 ----- ie both coins land with TT
   ii. P(Y = 1) = 1/2 ------ ie, either TH or HT
   iii. P(Y = 2) = 1/4 ------ ie, both coins with HH

   When you add all the probabilities we get 1, ie 1/4 + 1/2 + 1/4 =1

2. Consider the **rolling of a dice**. Let's say that the random variable, Y, is the number on the top face of a die when it is rolled once. The possible values for Y will thus be 1, 2, 3, 4, 5, and 6. The probability (P) of each of these values is 1/6 as they are all equally likely to be the value of Y

3. Consider the **rolling of three dice**. Let's say that the random variable Y may be the sum of the resulting numbers after three dice are rolled. Y could be 3 (1 + 1+ 1) or18 (6 + 6 + 6) or somewhere between 3 and 18 since the highest number of a die is 6 and the lowest number is 1.



In such types of uncertain processes the possible outcome is estimated using Monte Carlo method (also called as Monte Carlo simulation)

Monte Carlo method

Monte Carlo method is a mathematical technique used to estimate the possible outcomes of an uncertain events through repeated random sampling. It finds its application in many fields including business, physics, artificial intelligence, astronomy and engineering.

NOTE: The technique was initially developed by Stanislaw Ulam, a mathematician who worked on the Manhattan Project, the secret project on nuclear bomb during II world war

How does it works?

In this method, the repeated random sampling and recalculating the results again and again is very important. However each time use different set of random numbers between the minimum and

maximum values. This exercise can be repeated thousands of times (or even more) to obtain a large number of likely outcomes. Monte Carlo techniques involves four basic steps

1. Set up the *predictive model (domain identification)*, identifying both the dependent variable and the independent variables (also known as the input, risk or predictor variables)
2. Specify *probability distributions (generate unbiased inputs)*of the independent variables.
3. Run *simulations repeatedly (repeated random sampling)*, generating random values of the independent variables.
4. Do this until *enough results are gathered* to make up a representative sample of the near infinite number of possible combinations.

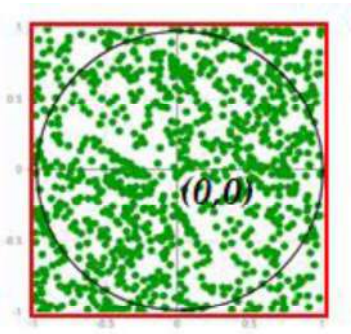
For ex, you want to find the average height of humans in India. The most accurate method is to find the height of country's entire population and take the average. Any way it is impossible. What is the next alternate? Find the height of few people randomly (say height of 100 members) and then take the average. If you do this only, say, in Karnataka the output (answer) is not a reliable one. Then what to do?? Follow these steps

i. Identify few states covering North, East, South and West part of the country(Identify the domains of possible inputs)
ii. Since height has inherent uncertainty, one has to measure the heights of selected people within the selected state (generate the inputs)
iii. The sample selection must remain unbiased. If surveyors collect samples of tall or short people, it will not give accurate results. Hence, the correct data would be obtained only through fair sample selection using a probability distribution.(generate the inputs randomly from a probability distribution over the domain)
iv. Take the average of the measured values.
v. Perform this procedure over and over, each time using a different set of random numbers between the minimum and maximum values (Repeated random sampling))
vi. As samples would represent many people, data collectors need to use more and more people as random samples. The greater the number of people, the better and more accurate range of results will be.(repetition of the exercise thousands of times to produce a better outcomes)

To determine the value of π

One of the basic examples of getting started with the Monte Carlo method is the estimation of π (pi). It is as follows

1. Consider a square of side length l = 2m and enclosing a circle of radius r = 1m. The center of both circle and square coincides at (0,0)
2. Generate a large number of uniformly distributed random points. These points can be in any position within the square i.e. between (0,0) and (1,1).
3. Keep track of the total number of points (N_T), and the number of points that are inside the circle (N_i)
4. We can show that the ratio of N_i and N_T is equal to the ratio of area of the circle and the square



$$\frac{N_i}{N_T} = \frac{area\ of\ the\ circle}{area\ of\ the\ square} = \frac{\pi r^2}{2^2}$$

$$\frac{N_i}{N_T} = \frac{\pi}{4}$$

$$\pi = 4 \times \frac{N_i}{N_T}$$

5. By substituting the values of N_i and N_T in the above equation we get the value of π
6. When we only have a small number of points, the estimation is not very accurate, but when we have hundreds of thousands of points, we get much closer to the actual value - to within around 2 decimal places of accuracy. It can be verified using following examples

i. If $N_i = 169$, & $N_T = 200$ then $\pi = 3.38$
ii. If $N_i = 468$, & $N_T = 600$ then $\pi = 3.12$
iii. If $N_i = 7067$, & $N_T = 9000$ then $\pi = 3.14$

Additional information:

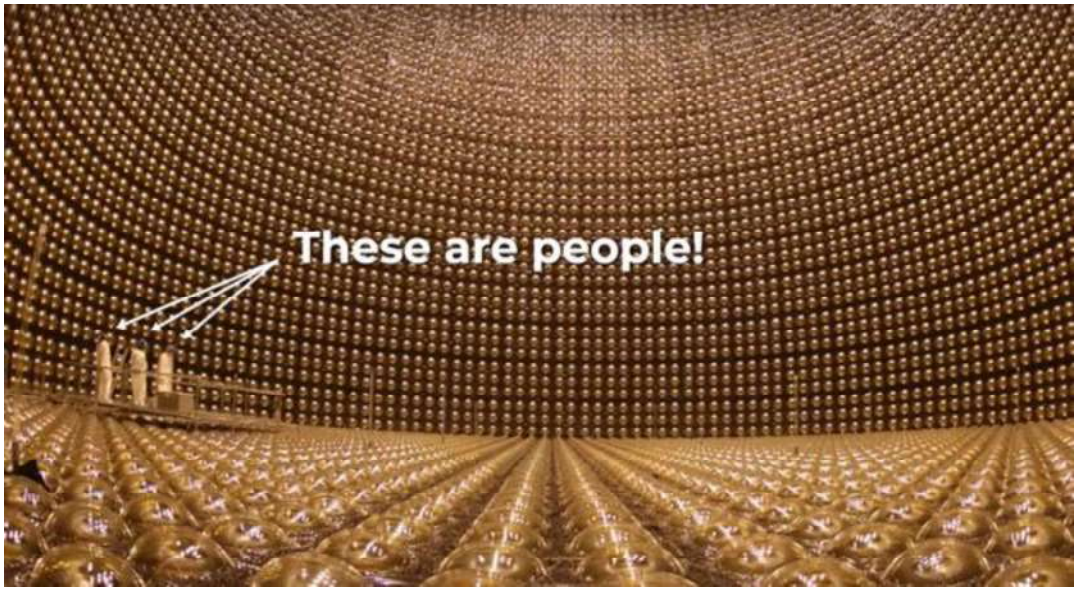


Image of Super-K plant

Some of the salient features of Super-K experiment in Japan

1. A cylindrical tank is buried 1km underground
2. The tank is 40 meter tall and 130ft radius
3. This tank contains 50,000 tons of ultra-pure water
4. 40% of the tank is covered by the detectors
5. In 1996 the scientists of super-K turned their detector ON and started looking for proton decay
6. Even today (after 27 years) not a single decay is detected!