Basic Quantum Physics

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF PHYSICS

UNIT – I - PHYSICS FOR ENGINEERS-SPHA1101

UNIT 1

BASIC QUANTUM PHYSICS

1.0 INTRODUCTION

Quantum theory is the theoretical basis of modern physics that explains the behaviour of matter and energy on the scale of atoms and subatomic particles / waves where classical physics does not always apply due to wave-particle duality and the uncertainty principle.

In 1900, physicist Max Planck presented his quantum theory to the German Physical Society. Planck had sought to discover the reason that radiation from a glowing body changes in color from red, to orange, and, finally, to blue as its temperature rises.

The Development of Quantum Theory

- In 1900, Planck made the assumption that energy was made of individual units, or quanta.
- In 1905, Albert Einstein theorized that not just the energy, but the radiation itself was *quantized*in the same manner.
- In 1924, Louis de Broglie proposed the wave nature of electrons and suggested that all matter has wave properties. This concept is known as the de Broglie hypothesis, an example of wave–particle duality, and forms a central part of the theory of quantum mechanics.
- In 1927, Werner Heisenberg proposed uncertainty principle. It states that the more precisely the position of some particle is determined, the less precisely its momentum can be known, and vice versa. The formal inequality relating

the standard deviation of position σ_x and the standard deviation of momentum σ_p was derived by Earle Hesse Kennard[3] later that year and by Hermann Weyl in 1928

$$\sigma_x \sigma_y \;\geq \frac{h}{4\pi}$$

where $\hbar$ is the reduced Planck constant, $h/(2\pi)$.

1.1 ELECTROMAGNETIC WAVES

Electric and magnetic fields in an electromagnetic wave are perpendicular to each other, *and* to the direction of propagation. Theexistence of electromagnetic waves, the only familiar electromagnetic waves was the visible light waves. The existence of ultraviolet and infrared waves was barely established. By the end of the nineteenth century, X-rays and gamma rays had also been discovered. We now know that, electromagnetic waves include visible light waves, X-rays, gamma rays, radio waves, and microwaves, ultraviolet and infrared waves. The classification of electromagneticwaves according to frequency is the electromagnetic spectrum. There is no sharp division between one kind of wave and the next. The classification is based roughly on how the waves are produced and/or detected.

The electric and magnetic fields $E_x$and $B_y$are perpendicular to each other, and to the direction z of propagation. We can write $E_x$and $B_y$as follows:

$$E_x = E_0 \sin(kz - \omega t)$$

$$B_y = B_0 \sin(kz - \omega t)$$

Here k is related to the wave length λ of the wave by the usual equation

$$k = \frac{2\pi}{\lambda}$$

Radio waves

Radio waves are produced by the accelerated motion of charges in conducting wires. They are used in radio and television communication systems. They are generally in the frequency range from 500 kHz to about 1000 MHz. The AM (amplitude modulated) band is from 530 kHz to 1710 kHz. Higher frequencies upto 54 MHz are used for *short wave* bands. TV waves range from 54 MHz to 890 MHz. The FM (frequency modulated) radio band extends from 88 MHz to 108 MHz. Cellular phones use radio waves to transmit voice communication in the ultrahigh frequency (UHF) band.

Microwaves

Microwaves (short-wavelength radio waves), with frequencies in the gigahertz (GHz) range, are produced by special vacuum tubes (calledklystrons, magnetrons and Gunn diodes). Due to their short wavelengths,they are suitable for the radar systems used in aircraft navigation. Radaralso provides the basis for the speed guns used to time fast balls, tennisserves,and automobiles. Microwave ovens are an interesting domesticapplication of these waves. In such ovens, the frequency of the microwavesis selected to match the resonant frequency of water molecules so thatenergy from the waves is transferred efficiently to the kinetic energy ofthe molecules. This raises the temperature of any food containing water.

Infrared waves

Infrared waves are produced by hot bodies and molecules. This band lies adjacent to the low-frequency or long-wave length end of the visible spectrum. Infrared waves are sometimes referred to as *heat waves*. This is because water molecules present in most materials readily absorb infrared waves (many other molecules, for example, CO_2, NH_3, also absorb infrared waves). After absorption, their thermal motion increases, that is, they heat up and heat their surroundings.

Infrared lamps are used in physical therapy. Infrared radiation also plays an important role in maintaining the earth's warmth or average temperature through the greenhouse effect. Incoming visible light (which passes relatively easily through the atmosphere) is absorbed by the earth's surface and reradiated as infrared (longer wavelength) radiations. This radiation is trapped by greenhouse gases such as carbon dioxide and water vapour. Infrared detectors are used in Earth satellites, both for military purposes and to observe growth of crops. Electronic devices (for example semiconductor light emitting diodes) also emit infrared and are widely used in the remote switches of household electronic systems such as TV sets, video recorders and hi-fi systems.

Visible rays

It is the most familiar form of electromagnetic waves. It is the part of the spectrum that is detected by the human eye. It runs from about 4×10^{14} Hz to about 7×10^{14} Hz or a wavelength range of about 700 – 400 nm. Visible light emitted or reflected from objects around us provides us information about the world. Our eyes are sensitive to this range of wavelengths. Different animals are sensitive to different range of wavelengths. For example, snakes can detect infrared waves, and the'visible' range of many insects extends well into the utraviolet.

Ultraviolet rays

It covers wavelengths ranging from about 4×10^{-7} m (400 nm) down to 6×10^{-10}m (0.6 nm). Ultraviolet (UV) radiation is produced by special lamps and very hot bodies. The sun is an important source of ultraviolet light. But fortunately, most of it is absorbed in the ozone layer in the atmosphere at an altitude of about 40 – 50 km. UV light in large quantities has harmful effects on humans. Exposure to UV radiation induces the production of more melanin, causing tanning of the skin. UV radiation is absorbed by ordinary glass. Hence, one cannot get tanned or sunburn through glass windows.

Welders wear special glass goggles or face masks with glass windows to protect their eyes from large amount of UV produced by welding arcs. Due to its shorter wavelengths, UV radiations can be focused into very narrow beams for high precision applications such as LASIK (*Laser assistedin situ keratomileusis*) eye surgery. UV lamps are used to kill germs in water purifiers.

Ozone layer in the atmosphere plays a protective role, and hence its depletion by Chlorofluorocarbons (CFCs) gas (such as Freon) is a matter of international concern.

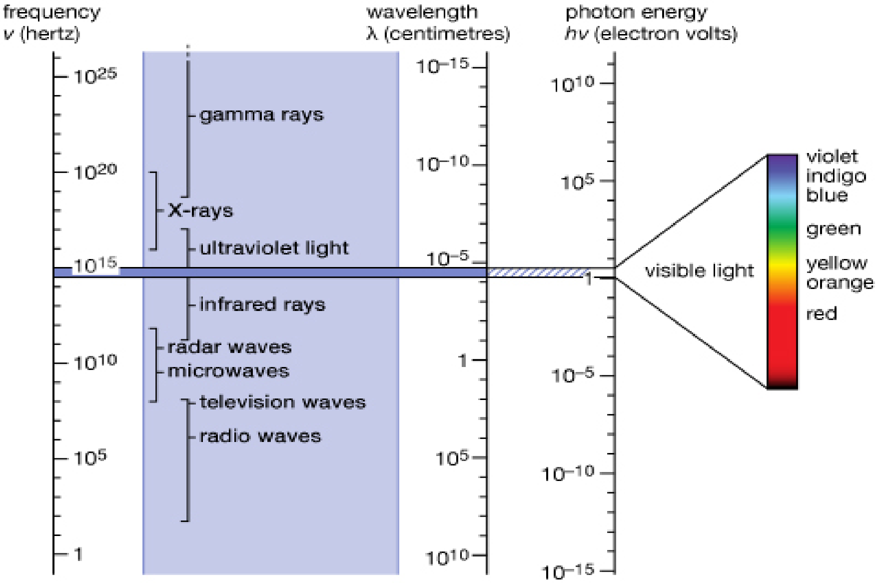


Fig.1.1 Electromagnetic Spectrum

X-rays

Beyond the UV region of the electromagnetic spectrum lies the X-ray region. We are familiar with X-rays because of its medical applications. It covers wavelengths from about 10^{-8} m (10 nm) down to 10^{-13} m (10^{-4} nm). One common way to generate X-rays is to bombard a metal target by high energy electrons. X-rays are used as a diagnostic tool in medicine and as a treatment for certain forms of cancer. Because X-rays damage or destroy living tissues and organisms, care must be taken to avoid unnecessary or over exposure.

Gamma rays

They lie in the upper frequency range of the electromagnetic spectrum and have wavelengths of from about 10^{-10}m to less than 10^{-14}m. This high frequency radiation is produced in nuclear reactions and also emitted by radioactive nuclei. They are used in medicine to destroy cancer cells.

1.2 PHOTOELECTRIC EFFECT

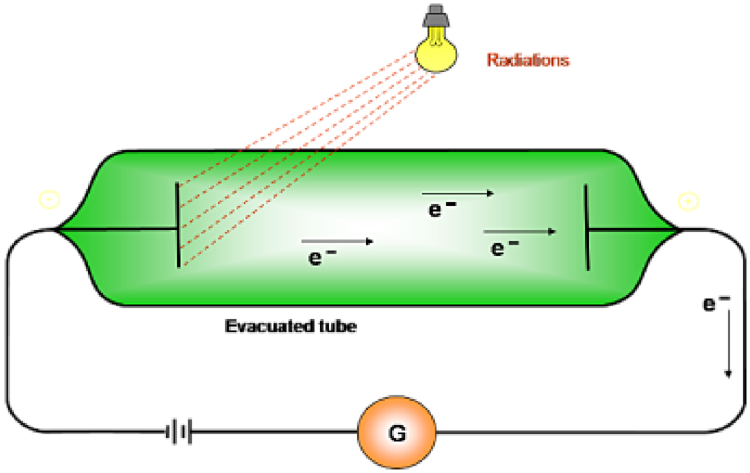


Fig.1.2

J.J. Thomson, observed that when a light of certain frequency strikes the surface of a metal, electrons are ejected. This phenomenon is known as photoelectric effect and the ejected electrons are called photoelectrons.

A few metals, which are having low ionization energy like Cesium, show this effect under the action of visible light but many more show it under the action of more energetic ultraviolet light.

An evacuated tube contains two electrodes connected to a source of variable voltage, with the metal plate whose surface is irradiated as the anode. Some of the photoelectrons that emerge from this surface have enough energy to reach the cathode despite its negative polarity, and they constitute the measured current. The slower photoelectrons are repelled before they get to the cathode. When the voltage is increased to a certain value V0, of the order of several volts, no more photoelectrons arrive, as indicated by the current dropping to zero.

This extinction voltage (or also referred as stopping potential) corresponds to the maximum photoelectron kinetic energy.

$$\text{i.e., } eV_o = \tfrac{1}{2} mv^2$$

The experimental findings are summarized as below:

- Electrons come out as soon as the light (of sufficient energy) strikes the metal surface.
- The light of any frequency will not be able to cause ejection of electrons from a metal surface. There is a minimum frequency, called the threshold (or critical) frequency, which can just cause the ejection. This frequency varies with the nature of the metal. The higher the frequency of the light, the more energy the photoelectrons have. Blue light results in faster electrons than red light.
- Photoelectric current is increased with increase in intensity of light of same frequency, if emission is permitted i.e., a bright light yields more photoelectrons than a dim one of the same frequency, but the electron energies remain the same.

Basic Quantum Physics

Light must have stream of energy particles or quanta of energy (hv). Suppose, the threshold frequency of light required to eject electrons from a metal is n0, when a photon of light of this frequency strikes a metal it imparts its entire energy (hv_0) to the electron.

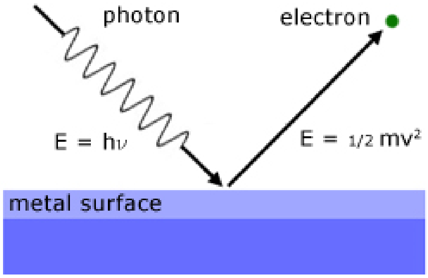


Fig.1.3

“This energy enables the electron to break away from the atom by overcoming the attractive influence of the nucleus”. Thus each photon can eject one electron. If the frequency of light is less than v_0 there is no ejection of electron. If the frequency of light is higher than v_0 (let it be v), the photon of this light having higher energy (hv), will impart some energy to the electron that is needed to remove it from the atom. The excess energy would give a certain velocity (i.e, kinetic energy) to the electron.

$$hv = hv_0 + K.E$$

$$hv = hv_0 + \frac{1}{2} mv^2$$

$$\frac{1}{2} mv^2 = hv - hv_o$$

Where, **v = frequency of the incident light**

v_0 = threshold frequency

$h\nu_0$ is the threshold energy (or) the work function denoted by $\Phi = h\nu_0$ (minimum energy of the photon to liberate electron). It is constant for particular metal and is also equal to the ionization potential of gaseous atoms.

The kinetic energy of the photoelectrons increases linearly with the frequency of incident light. Thus, if the energy of the ejected electrons is plotted as a function of frequency, it result in a straight line whose slope is equal to Planck's constant 'h' and whose intercept is hn_0.

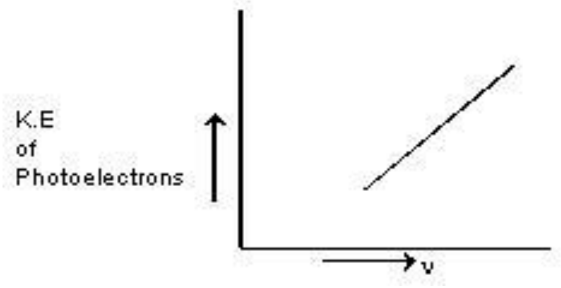


Fig.1.4

1.3 COMPTON SCATTERING

Photoelectric effect provides evidence that energy is quantized. In order to establish the particle nature of radiation, it is necessary that photons must carry momentum. In 1922, Arthur Compton studied the scattering of x-rays of known frequency from graphite and looked at the recoil electrons and the scattered x-rays. According to wave theory, when an electromagnetic wave of frequency is incident on an atom, it would cause electrons to oscillate. The electrons would absorb energy from the wave and re-radiate electromagnetic wave of a frequency$v_s > v_o$. The frequency of scattered radiation would depend on the amount of energy absorbed from the wave, i.e. on the intensity of incident radiation and the duration of the exposure of electrons to the radiation and not on the frequency of the incident radiation. Compton found that the wavelength of the scattered radiation does not depend on the intensity of incident radiation but it depends on the angle of scattering and the wavelength of the incident beam. The wavelength of the radiation scattered at an angle $\theta$is given by

$$\lambda_x = \lambda_0 + \frac{h}{\mathrm{m_0C}}(1 - Cos\theta)$$

Where m_0 is the rest mass of the electron. The constant $\frac{h}{m_0\mathrm{C}}$ is known as the Compton wavelength of the electron and it has a value 0.0024nm.

The spectrum of radiation at an angle θ consists of two peaks, one at λ_0 and the other at λ_x. The Compton effect can be explained by assuming that the incoming radiation is a beam of particles with

The energy

$$E = h\upsilon_0$$

The momentum

$$p = \frac{h}{m_0c}$$

For energy momentum relation

$$E^2 = p^2c^2 + mc^2$$

1.4 PHOTON

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

Photons are always in motion and, in a vacuum, travel at a constant speed to all observers of 2.998 × 10^8m/s. This is commonly referred to as the speed of light, denoted by the letter *c*.

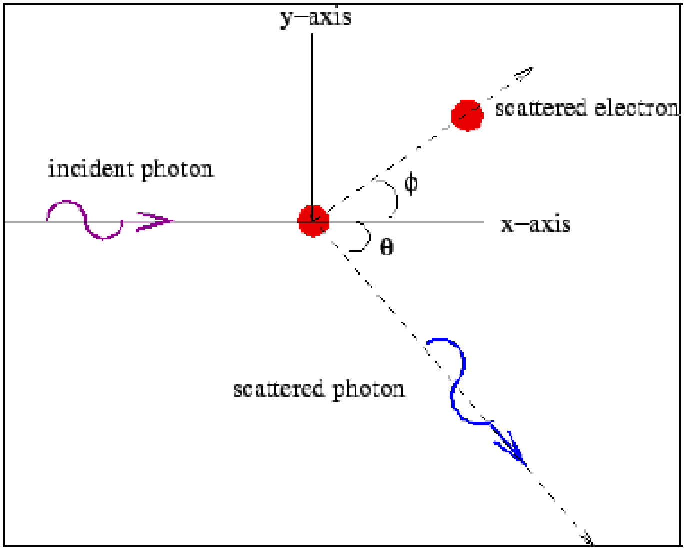


Fig.1.6

As per Einstein's light quantum theory, photons have energy equal to their oscillation frequency times Planck's constant. Einstein proved that light is a flow of photons, the energy of these photons is the height of their oscillation frequency, and the intensity of the light corresponds to the number of photons. Essentially, he explained how a stream of photons can act both as a wave and particle.

Photon properties

The basic properties of photons are:

- Zero mass and rest energy and exist as moving particles.
- Elementary particles despite lacking rest mass.

- No electric charge.
- Stable.
- Carry energy and momentum which are dependent on the frequency.
- Have interactions with other particles such as electrons, such as the Compton effect.
- They can be destroyed or created by many natural processes, for instance when radiation is absorbed or emitted.
- When in empty space, they travel at the speed of light.

1.5 FRANCK-HERTZ EXPERIMENT

James Franck and Gustav Hertz conducted an experiment in 1914, which demonstrated the existence of excited states in mercury atoms. It confirms the prediction of quantum theory that electrons occupy only discrete, quantized energy states. This experiment supports Bohr model of atom. For this great invention they have been awarded Nobel Prize in Physics in the year 1925.

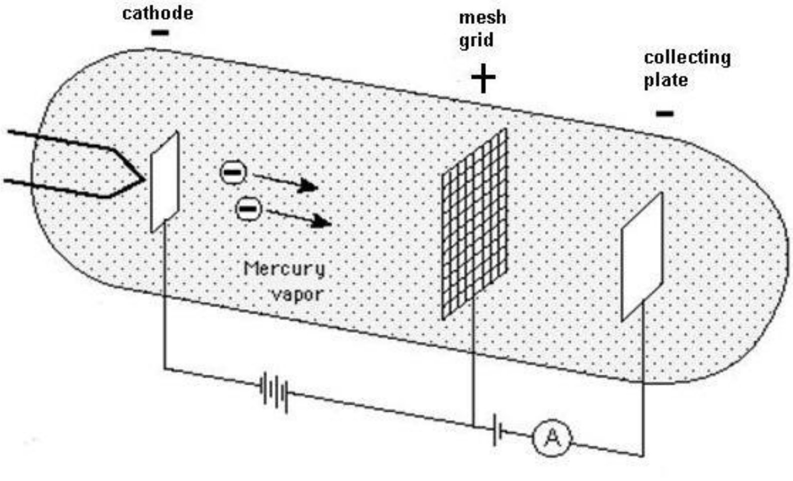


Fig.1.7 Experimental setup

With the help of thermionic emission, electrons are emitted by a heated cathode, and then accelerated toward a grid which is at a positive potential, relative to the cathode. The collecting plate is at a lower potential and is negative with respect to mesh grid. If electrons have sufficient energy on reaching the grid, some will pass through the grid, and reach collecting plate, and it will be measured as current by the ammeter. Electrons which do not have sufficient energy on reaching the grid will be slowed down, and will fall back to the grid.The experimental results confirm the existence of discrete energy levels.

As long as the electron collision is elastic, the electrons will not lose energy on colliding with gas molecules in tube. As the accelerating potential increases, the current also increases. But as the accelerating potential reaches a particular value, (4.9eV for mercury, 19eV for neon), each electron possess that much of potential and now the collision become inelastic. As a result, the energy level of electron bound to the atom is raised. Now the electron almost loses its energy, and measured current drops.

1.6 BOHR ATOM

As we know it the atom is made up of neutrons, protons, and electrons. While the neutrons (no charge), and the protons (positive charge) make up the nucleus of the atom, the electron (negative charge) circle around the nucleus much like the orbiting of planets around the sun. According to Rutherford Model electrons are in constant motion around the nucleus while emitting energy. According to this idea as the electrons lost energy they would get closer and closer to the nucleus and eventually crash into it. This problem was solved when Niels Bohr created a model of a hydrogen atom. Bohr decided to formulate an atom that would correct the instability problems of the Rutherford Model.

1. Electrons move around the nucleus of an atom in circular motion.

2. Electrons have a set number of orbitals, ring-like paths around the nucleus, they can travel in called *stationary states*. If the electron stays in one orbital, the energy of the electron remains constant. The first orbital is n=1, the second orbital n=2 and so on. (These values of n are called quantum numbers.)
3. Electrons can only move from one orbital to another allowed orbital at one time. If an electron drops down from n=2 to n=1 energy is emitted, and if the electron moves up from n=1 to n=2 energy is absorbed. The amount of energy that is either absorbed or emitted is called quanta.
4. The energy of a photon, lost or gained, is calculated using Planck's equation: (h is Planck's constant, 6.62607×10^{-34} J s / cycles, and v stands for frequency in cycles/s)

$$\text{Energy Difference } (\Delta E) = hv$$

Bohr's model allows us to calculate the radii of the orbits that are allowed for an electron to travel. It also allows us to calculate electron velocities and energy in these orbits. With the equation below we can calculate allowed energy levels for a Hydrogen atom. ($-R_H$ is a constant value of 2.179×10^{-18} Joules)

$$E_n = -R_H \frac{1}{n^2}$$

When an electron drops from a higher level for example from n=2 to a lower energy level, n=1, it emits a specific wavelength distinct to that element. To calculate the difference between the two energy levels, the equation below is used. Subscript i stands for initial, (initial energy= E_i and initial quantum number= n_i) while subscript f stands for final (final energy= E_f and final quantum number= n_f).

$$\text{Energy Difference } (\Delta E) = E_f - E_i = -R_H\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) = R_H\left(\frac{1}{n_i^2} - \frac{1}{n_f^2}\right)$$

The Bohr model also helps us understand how cations are made. This is through the concept of Ionization Energy, the energy to remove an electron from its ground state. After the electron is free the atom is ionized. E_i stands for the ionization energy of the hydrogen atom. The equation below however works for any species that have one electron like hydrogen (for example, He^+ and Li^{2+}) where Z is the atomic number of the hydrogen-like atom:

$$E_n = -Z^2 R_H \frac{1}{n^2}$$

If we want to find the energy associated with a change in quantum number of a hydrogen-like atom, we can combine this equation with the previous one to form

$$E_n = -Z^2 R_H \left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

1.7 ELECTRON DIFFRACTION

The wave-particle duality of nature results in particles having both wave-like and particle-like properties. The wave-like properties of objects means that we can perform diffraction using them, as with X-ray photons. However, this time we extend from massless particles to those with mass, including the electron and the neutron.

In this case, the wavelength of an electron is given by the de Broglie relation, as shown in Equation

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

where h is Planck's constant and p is the electron momentum. As the momentum is related to the velocity and, in turn, the velocity is dependent upon the accelerating force i.e. an electric potential we see that equationcan be written

$$\lambda = \frac{h}{\sqrt{2m_0 eU}}$$

where m_0 is the mass of the electron and U is the electric potential through which the electron is accelerated.

1.8 WAVE-PARTICLE DUALITY OF RADIATION AND MATTER

The dual nature of matter and the dual nature of radiation were revolutionary concepts of physics. In case of light some phenomenon like diffraction and interference can be explained on the basis of its wave character. However, the certain other phenomenon such as black body radiation and photoelectric effect can be explained only on the basis of its particle nature. Thus, light is said to have a dual character. Such studies on light were made by Einstein in 1905. Louis de Broglie, in 1924 extended the idea of photons to material particles such as electron and he proposed that matter also has a dual character-as wave and as particle.

S. No.	Radiation	Material
1	Radiation spreads in the form of waves.	Material shoots in straight line.
2	It has no mass.	Material has mass.
3	Radiation can only be experienced.	Material is Visible, it has volume etc.
4	Radiation has energy known as radiation energy.	It has energy known as mechanical energy.
5	Main representative of radiation is light.	Main representative of material is electron.

1.8.1 Derivation of de-Broglie Equation

The wavelength of the wave associated with any material particle was calculated by analogy with photon.

Basic Quantum Physics

In case of photon, if it is assumed to have wave character, its energy is given by

$$E = hv \quad \ldots(i)$$

(According to the Planck's quantum theory)

Where ν frequency of the wave and 'h' is is Planck's constant

If the photon is supposed to have particle character, its energy is given by

$$E = mc^2 \quad \ldots\ldots\ (ii)$$

(according to Einstein's equation)

where 'm' is the mass of photon, 'c' is the velocity of light.

By equating (i) and (ii)

$$hv = mc^2$$

$$\text{But } v = c/\lambda$$

$$h\,c/\lambda \;= mc^2$$

$$\text{(or) } \lambda = h/mc$$

The above equation is applicable to material particle if the mass and velocity of photon is replaced by the mass and velocity of material particle. Thus for any material particle like electron.

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

where mv = p is the momentum of the particle.

1.8.2 De-Broglie wave

De-Broglie first proposed that the light has both wave like and particle like properties. Electrons can also have wave like properties. The De Broglie equation is given by,

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

Here, λ is the de-Broglie wavelength, h is the Planck's constant, and p is the momentum.

For heavy particles which are having momentum, and can create a relation with the plank's constant, those particles are associated the de Broglie wavelength.

This relationship between the wavelength and momentum holds good for all types of matter, i.e. all matters exhibit the particle and wave nature properties.

De Broglie states that any particle should exhibit wave like nature and also particle like nature. The particle velocity should be always equal the group velocity of the corresponding wave.

The wave like nature of light is having two supportive experimental proofs as suggested in the De Broglie hypothesis. The experimental proofs are

- The discrete atomic energy levels of the atom in the electronic structure.
- And the diffraction pattern of electrons from the crystal planes in solid materials. The atomic energy levels, the electron waves can be analyzed as a constructive interference pattern by using the Bohr model.
- Again, the wave nature of the electrons can be explained, by the particle in box analysis in quantum mechanics. In this case, to explain the behavior of the electron wave, we can use a box in which the electron is confined, and is having the dimensions in the order of the size of an atom.

Physical interpretation of a wave function

A wave function of quantum mechanical system is some thing associated with its wave nature. To arrive at results consistent with real timeobservations, we have imposed several requirements on the wave function of a quantum mechanical system. Wave function is a quantity which describe the particle.**Wave function must be single valued and continuous everywhere** and it should **approach to zero as r is tending to infinity.**

It is a mathematical function that describes some of the properties of the wave. the shape of the wave is described by the sine function, and the amplitude describes strength (intensity) of the wave. The wave function can be interpreted as **ψ(r,t).**

$$\psi = A\sin\theta$$

The probability of finding a particle at a position$x_1 < x > x_2$and time t

$$\int_{x1}^{x2} |\Psi(x)|^2 dx$$

OR

$$\int_{x1}^{x2} \psi^* \psi dx$$

The probability per unit length of finding the particle at the position **r** at time **t**.**ψ*(r,t)** is complex conjugate of **ψ(r,t)**

Requirements for An Acceptable Wavefunction

1. The wave function **ψmust be continuous**. All its **partial derivatives must also becontinuous** (partial derivatives are$\left(\frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y}\text{etc.}\right)$ This makes the wave function"smooth".
2. The wave function ψ must be **quadratically integrable**. This means that the integral $\int \psi^* \psi d\tau$must exist.

3. Since $\int\psi^* \psi d\tau$ is the probability density, **it must be single valued**.
4. The wave functions must form an **orthonormal set**. This means that
    - the wave functions must be **normalized.**

      $\int_{-\infty}^{\infty} \psi_i * \psi_i \mathrm{d}\tau = 1$
    - the wave functions must be **orthogonal**.

      $\int_{-\infty}^{\infty} \psi_j * \psi_j \mathrm{d}\tau=0$

$\int_{-\infty}^{\infty} \psi_j * \psi_j \mathrm{d}\tau =\delta_{ij}$ where$\delta_{ij} = 1\, wheni = j;\ \delta_{ij} = 0\, wheni \neq j$

Δijis called Kronecker delta

5. The wave function must be **finite everywhere**.
6. The wave function must satisfy the **boundary conditions** of the quantum mechanical system it represents.

1.9 HEISENBERG UNCERTAINTY PRINCIPLE

Uncertainty principle, also called **Heisenberg uncertainty principle** or **indeterminacy principle**, statement, articulated (1927) by the German physicist Werner Heisenberg, that the position and the velocity of an object cannot both be measured exactly, at the same time, even in theory. The very concepts of exact position and exact velocity together, in fact, have no meaning in nature.

The uncertainty principle is alternatively expressed in terms of a particle's momentum and position. The momentum of a particle is equal to the product of its mass times its velocity. Thus, the product of the uncertainties in the momentum and the position of a particle equals $h/(4\pi)$ or more, where h is Planck's constant, or about 6.6×10^{-34} joule-second). It is impossible to know the precise position and momentum. This relationship also applies to energy and time, in that one cannot measure the precise energy of a system in a finite amount of time. Uncertainties in the products of "conjugate pairs" (momentum/position) and (energy/time) were defined by Heisenberg as having a minimum value corresponding to Planck's constant divided by 4π. More clearly:

$$\Delta p * \Delta x \geq \frac{h}{4\pi}$$

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

Where Δ refers to the uncertainty in that variable.

Qualitatively this principle states that "the order of magnitude of the product of the uncertainties in the knowledge of two variables must be at least Planck's constant h". Considering the position and momentum is the pair of physical variables, we have

$$\Delta p\ \Delta x\ \approx h \qquad \text{... (1)}$$

where Δp is the uncertainty in determining the momentum and

Δx is the uncertainty in determining the position of the particle.

Similarly, we have

$$\Delta E\ \Delta t\ \approx h \qquad \text{... (2)}$$

$$\Delta \mathrm{J}\ \Delta\theta\ \approx \mathrm{h} \qquad \text{... (3)}$$

where ΔE and Δt are uncertainties in determining the energy and time while ΔJ and Δθ uncertainties in determining the angular momentum and angle.

Applications of Uncertainty Principle

1.9.1 Calculation of the radius of Bohr's first orbit

Considering Δx and Δp are the uncertainties in position and momentum of the electron in the first orbit, then according to uncertainty principle,

$$(\Delta x)(\Delta p) \geq \frac{h}{4\pi}$$

$$(\Delta p) \geq \frac{h}{(\Delta x)\, 2\pi} \qquad \text{using } (\Delta p)\,(\Delta x) \approx \frac{h}{2\pi}$$

The uncertainty in the kinetic energy (ΔT) of the electron may be written as:

$$\Delta T \geq \frac{1}{2} m\, (\Delta \nu)^2 \geq \frac{m^2 (\Delta \nu)^2}{2m} \geq \frac{(\Delta p)^2}{2m}$$

$$\Delta T \geq \left[\frac{1}{2m}\right] \frac{h^2}{4\pi^2 (\Delta x)^2}$$

The uncertainty in the potential energy ΔV of the same electron is given by:

$$\Delta V \geq -\frac{Ze^2}{4\pi\, \epsilon_0 (\Delta x)}$$

Hence the uncertainty in the total energy ΔE is given by:

$$\Delta E \;\geq\; \Delta T + \;\Delta V \;\geq\; \frac{h^2}{4\pi^2 \times 2m\, (\Delta x)^2} - \frac{Ze^2}{(\Delta x) 4\pi\epsilon_0}$$

The uncertainty in the energy will be minimum, if

$$\frac{d(\Delta E)}{d(\Delta x)} = 0 \quad \text{and} \quad \frac{d^2(\Delta E)}{d^2(\Delta x)} = \text{positive}$$

Now

$$\frac{d(\Delta E)}{d(\Delta x)} \geq \frac{-2h^2}{8\pi^2\, m\, (\Delta x)^3} + \frac{Ze^2}{4\pi\epsilon_0 (\Delta x)^2} = 0$$

$$\text{i.e.} \qquad \frac{h^2}{4\pi^2\, m\, (\Delta x)^3} = \frac{Ze^2}{4\pi\epsilon_0 (\Delta x)^2}$$

$$\Delta x = \frac{h^2 \epsilon_0}{\pi m\, Ze^2}$$

Thus the radius of the first orbit of hydrogen is:

$$r \approx \Delta x \approx \frac{h^2 \epsilon_0}{\pi m\, e^2} \qquad \text{with Z = 1 and n = 1}$$

$$r = \frac{h^2 \epsilon_0}{\pi m\, e^2}$$

1.9.2 Calculation of energy of particle in 1D box:

Consider a particle of mass ‘m’ moving along the x-axis between the two rigid walls of the well with x = 0 and x = a.

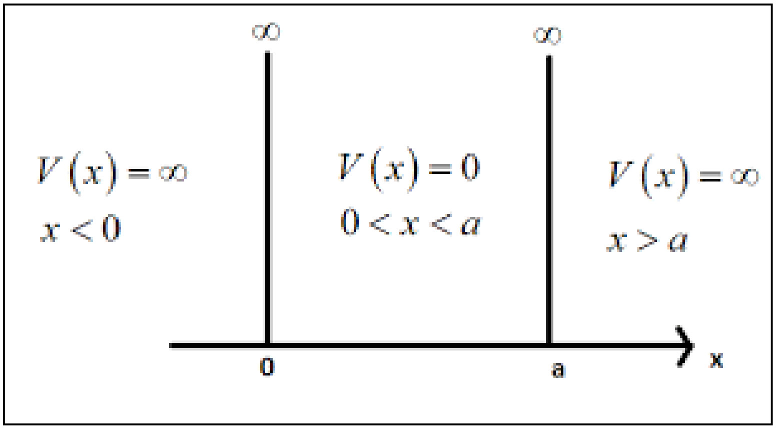


Fig.1.8 Particle in a box

When the force acting on the particle between the walls is zero, its potential energy is constant in this region and it is duly taken to be zero. As the walls are assumed to be rigid, the force acting on the particle abruptly increases from zero to finite value at the boundaries and hence the potential energy of the particle becomes infinitely large at x = 0 and x = a.

Thus, potential function V(x) is expressed as

$$V(x) = 0 \text{ for } 0 < x < a$$

$$V(x) = \infty \text{ for } 0 \geq x \geq a$$

This potential function is known as square well potential.

The particle cannot move out of the box. Also, it cannot exist on the walls of the box. So its wave function Ψ is 0 for $x \leq 0$ and $x \geq a$ and for zero potential the Schrodinger equation can be written as:

$$\frac{d^2\Psi}{dx^2} + \frac{2mE}{\hbar^2}\Psi = 0 \qquad \text{... (1)}$$

$$\frac{d^2\Psi}{dx^2} + k^2\Psi = 0 \qquad \text{... (2)}$$

where $k^2 = \frac{2mE}{\hbar^2}$

The general solution of equation (2) is

$$\Psi(x) = A\, sinkx + B \cos kx \qquad \text{... (3)}$$

Applying the boundary conditions $\Psi(x) = 0$ at x = a and x = 0, one gets $k = \frac{n\pi}{a}$ where n = 1, 2, 3, ... and B = 0.

Thus the wave function associated with the electron is:

$$\Psi_n = A \sin\left[\frac{n\pi x}{a}\right] \qquad \text{... (4)}$$

$$E_n = \frac{k^2\hbar^2}{2m} = \frac{n^2h^2}{8ma^2} \qquad \text{... (5)}$$

This equation gives the energy of the particle in the n^{th} energy state.

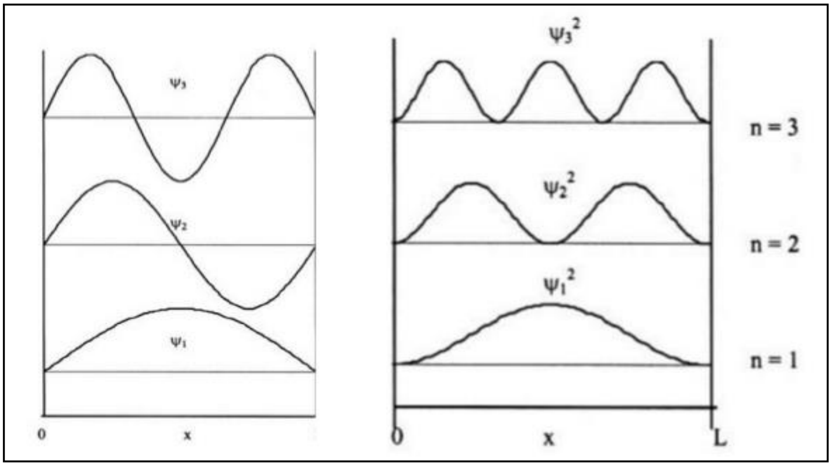


Fig.1.9 Energy levels and wave functions

The particle in the box cannot possess any arbitrary amount of energy. Rather, it can have discrete energy values specified by the equation (5). In other words, its energy is quantised.

Each value of E_n is called as eigen value and the corresponding Ψ_n is called eigen function.

1.10 OPERATORS

Operators are defined as any mathematical operation, differentiation, integration, division, multiplication, addition, subtraction etc. can be represented by certain symbols. In another words an operator O is a mathematical operation which may be applied to a function f(x) which changes the function to another function g(x). This can be represented as

$$O\,f(x) = g(x)$$

For example, $\frac{d}{dx}(x^3 + 1) = 3\,x^2$

In operator language, it may be stated that when the operator $\frac{d}{dx}$ operates on function $(x^3 + 1)$, it changes to $3\,x^2$.

Now we shall illustrate the importance of remembering the order in which the individual operations must be carried out.

For example, if we first operate function f(x) on the operator x, getting x f(x), and next operate the result with $\frac{d}{dx}$ then the final result is

$$\frac{d}{dx}[xf(x)] = x\,\frac{d}{dx}\,f(x) + f(x)$$

Now if we operate f(x) first with $\frac{d}{dx}$ and then the result with operator x, then the result is $x\,\frac{d}{dx}\,f(x)$

so that

$$\frac{d}{dx}[xf(x)] \;\neq\; x\,\frac{d}{dx}\,f(x)$$

Eigenvalues and Eigenfunctions:

If there is a class of functions which, when operated by an operator α, are merely multiplied by some constant λ, i.e., if

$$\alpha f(x) = \lambda\, \mathrm{f(x)}$$

then the functions f(x) are called eigenfunctions of the operator α and the various possible values λ are called eigenvalues of the operator.

For example, let us consider the function sin 4π. If it is operated by an operator $\frac{-d^2}{dx^2}$, the result is

$$\frac{-d^2}{dx^2}(\sin 4x) = 16 \sin 4x$$

Thus sin 4x is eigenfunction and 16 is the eigenvalue of the operator $\frac{-d^2}{dx^2}$. In quantum mechanics, the allowed eigenfunctions are continuous, finite and single valued.

1.11 SCHRÖDINGER WAVE EQUATION

Schrodinger wave equation describes the wave nature of a particle in mathematical form. It is the basis equation of wave mechanics originally proposed by Austrian Scientist Erwin Schrödinger.

Schroedinger connected the expression of De-Broglie wavelength with the classical wave equation for a moving particle and obtained a new wave equation.

Forms of Schoredinger wave equation

There are two forms of Schroedinger wave equation. They are

i) Time independent wave equation

ii) Time dependent wave equation

1.11.1 Schroedinger's time independent wave equation (Stationary states)

According to De-Broglie's hypothesis, a particle of mass m moving with velocity v is always associated with a wave whose wavelength is given by $\lambda = \frac{h}{mv}$. If the particle has wave properties, then it is expected that there should be some sort of wave equation to describe the behaviour of the particle.

The classical differential equation of wave system is

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2 \Psi}{\partial t^2} \qquad \text{... (1)}$$

where Ψ is wave displacement for De-Broglie’s waves at time t;

x,y,z are coordinates of the particle ; and v is wave velocity

Equation (1) may be rewritten as

$$\nabla^2\Psi = \frac{1}{v^2}\frac{\partial^2\Psi}{\partial t^2} \qquad \text{... (2)}$$

where $\nabla^2\Psi = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is Laplacian operator.

The solution of equation (2) gives the periodic displacement in terms of time t,

i.e. $\Psi\,(x,y,z,t) = \Psi_0(x,y,z)e^{-i\omega t}$... (3)

where Ψ_0 is amplitude at the point considered . It is a function of position and not of time (t).

Differentiating equation (3) twice with respect to t, we get

$$\frac{\partial\Psi}{\partial t} = -i\,\omega\,\Psi_0 e^{-i\omega t}$$

and $\frac{\partial^2\Psi}{\partial t^2} = (-i\omega)(-i\omega)\Psi_0 e^{i\omega t} = i^2\omega^2\Psi_0 e^{-i\omega t}$... (4)

$$= -\omega^2\,\Psi \qquad [\,i^2 = -1]$$

Substituting this in equation (2), we get

$$\nabla^2\Psi = -\frac{\omega^2}{v^2}\,\Psi \qquad \text{... (5)}$$

or $\nabla^2\Psi + \frac{4\pi^2}{\lambda^2}\,\Psi = 0$... (6)

$\nabla^2\Psi + \frac{4\pi^2 m^2 v^2}{h^2}\,\Psi = 0$ ... (7)

If E is total energy of the particle, V is potential energy and $\frac{1}{2}mv^2$ is kinetic energy, then

$$E = P.E. + K.E.$$

i.e., $E = V + \frac{1}{2}mv^2$ or $E - V = \frac{1}{2}mv^2$

$$mv^2 = 2(E - V)$$

Multiply m on both sides

or $m^2v^2 = 2m\,(E - V)$ ... (8)

Substituting equation (8) in equation (7) , we get

$$\nabla^2\Psi + \frac{8\pi^2 m}{h^2}(E - V)\,\Psi = 0 \qquad \text{... (9)}$$

This Ψ is independent of time. This is the steady state (stationary state) form of Schroedinger equation. The equation is known as **Schroedinger's time independent wave equation.**

If we consider one-dimensional motion, i.e., particle moving along x-direction only, then Schroedinger's equation (9) reduces to

$$\frac{d^2\Psi(x)}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V)\Psi(x) = 0 \qquad \text{... (10)}$$

Taking $\hbar = \frac{h}{2\pi}$ (where ħ is a reduced Planck's constant) in equation (9), Schroedinger's wave equation may be rewritten as

$$\nabla^2\Psi + \frac{2m}{\hbar^2}(E - V)\,\Psi = 0 \qquad \text{... (11)}$$

1.11.2 Schroedinger's time dependent wave equation:

We can easily derive Schroedinger's time dependent wave equation by eliminating 'E' from Schroedinger's time independent wave equation.

Differentiating equation (3) with respect to time t, we get

$$\frac{\partial \Psi}{\partial t} = -i\,\omega\,\Psi_0 e^{-i\omega t} = -i\,(2\pi\nu)\Psi_0 e^{-i\omega t} \qquad (\omega = 2\pi\nu)$$

$$= -2\pi\, i\,\nu\,\Psi = -i\,2\pi \frac{E}{h}\,\Psi \qquad (\nu = \frac{E}{h})$$

$$\therefore\ \frac{\partial \Psi}{\partial t} = -i\,\frac{E}{\frac{h}{2\pi}}\,\Psi = -i\,\frac{E}{\hbar}\,\Psi$$

Multiplying i on both sides,

$$i\frac{\partial \Psi}{\partial t} = -i\ \times i\ \left[\frac{E}{\hbar}\right]\Psi$$

$$i\frac{\partial \Psi}{\partial t} = \frac{E}{\hbar}\,\Psi$$

or $$E\,\Psi = i\,\hbar\,\frac{\partial \Psi}{\partial t} \qquad \text{... (12)}$$

Substituting the value of $E\,\Psi$ from eq. (12) in Schroedinger's time independent wave equation (11), we have

$$\nabla^2\Psi + \frac{2m}{\hbar^2}(E - V)\,\Psi = 0$$

$$\nabla^2\Psi + \frac{2m}{\hbar^2}(E\Psi - V\Psi) = 0$$

$$\nabla^2\Psi + \frac{2m}{\hbar^2}(i\,\hbar\,\frac{\partial \Psi}{\partial t} - V\Psi) = 0$$

$$\nabla^2\Psi = \frac{-2m}{\hbar^2}\left(i\,\hbar\,\frac{\partial\Psi}{\partial t} - V\Psi\right)$$

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\,\hbar\,\frac{\partial\Psi}{\partial t} \qquad \text{... (13)}$$

This equation is known as derive Schrodinger's time dependent wave equation.

1.12 PHYSICAL APPLICATIONS OF SCHRODINGER'S EQUATION TO SQUARE WELL POTENTIAL IN ONE DIMENSION: TRANSMISSION AND REFLECTION COEFFICIENT AT A BARRIER:

A potential barrier corresponds to the force field acting on a particle being zero everywhere except in a limited region. A single potential barrier is shown in Fig.1.

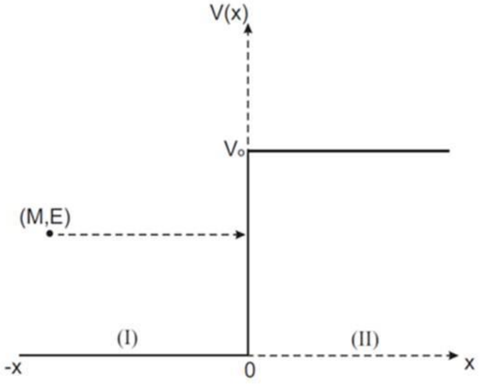


Fig.1.10

For x < 0, the potential V(x) = 0

For x > 0, the potential V(x) = V_0

So we may define the potential function of a potential step as Fig.1.10.

V(x) = 0 for x < 0

$V_0$for x > 0

$V_0$is considered as the height of the potential barrier. Now consider the case of electrons carrying energy E moving from left to right (i.e.) along the positive X-directing. The energy of the electron may be greater than V_0or less than V_0. According to quantum mechanics, the electrons behaves like a wave and while moving from left to right, it faces a sudden shift in potential at x = 0. This is analogous to the propagation of light wave when it strikes a glass plate and faces a change in refractive index. The ray is partly reflected and partly transmitted. In a similar way, the electrons will be partly reflected and partly transmitted at x = 0 i.e., at the discontinuity.

The Schrodinger wave equation is applied to region I and II at the condition E > V_0.

In region I, the Schrodinger wave equation is given by

$$\frac{d^2\Psi_1}{dx^2} + \frac{2mE}{\hbar^2}\Psi_1 = 0$$

it may be written as

or $\frac{d^2\Psi_1}{dx^2} k_1^2\Psi_1$ where $\frac{2mE}{\hbar^2} = k_1^2$ ... (1)

In region II, the Schrodinger wave equation is given by

$$\frac{d^2\Psi_2}{dx^2} + \frac{2mE}{\hbar^2}(E - V_0)\Psi_2 = 0$$

or $\frac{d^2\Psi_2}{dx^2} k_2^2\Psi_2 = 0$ where $\frac{2m(E-V)}{\hbar^2} = k_2^2$ ... (2)

The general solutions of equations (1) and (2) takes the form

$$\Psi_1 = A\,e^{ik_1x} + Be^{-ik_1x} \quad (x < 0) \qquad \text{... (3)}$$

and $$\Psi_2 = C\,e^{ik_2x} + De^{-ik_2x} \quad (x > 0) \qquad \text{... (4)}$$

Where A, B, C and D are constants which may be found boundary conditions.

In equation (3), the first term corresponds to a wave propagating along +ve X-direction while the second term corresponds to wave travelling along negative X-direction in the region I. While the first will represent incident wave, the second will represent reflected wave. Since in the first region both incident and reflected rays will be present, so equation (3) represents the solution of Schrodinger’s equation in region I.

In equation (4), the first term corresponds to a wave travelling along positive X-direction i.e., transmitted wave and second term corresponds to the wave travelling in negative X-direction i.e., reflected wave in region II. Since there is no discontinuity in region II, so there will not be a reflected wave i.e., D = 0, and so we have the solution of wave equation as

$$\Psi_2 = C\,e^{ikx} \quad \text{... (5)}$$

So equations (3) and (5) will represent the solutions of Schrodinger wave equation in I and II region respectively.

From probability interpretation of wave function, ψ must be finite i.e., it must be continuous. So we will have the following boundary conditions.

$$(\Psi_1)_{x=0} = (\Psi_2)_{x=0}$$

And $$\left(\frac{d\Psi_1}{dx}\right)_{x=0} = \left(\frac{d\Psi_2}{dx}\right)_{x=0} \qquad \text{... (6)}$$

Applying these boundary conditions, to equations (3) and (5), we get

$$A + B = 0 \qquad \ldots (7)$$

and $$k_1A - k_1B = k_2C \qquad \ldots (8)$$

Putting the value of C in equation (8), we have

$$k_1A - k_1B = k_2(A + B)$$

$$(k_1 - k_2)A = (k_1 + k_2)B$$

Which gives $$B = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)A \qquad \ldots (9)$$

Putting the value of B from equation (9) in equation (7), we get

$$A + \frac{(k_1 - k_2)}{(k_1 + k_2)}A = C$$

Which gives $$C = \left(\frac{2k_1}{k_1 + k_2}\right)A \qquad \ldots (10)$$

In this case B and C represent the amplitudes of reflected and transmitted beam respectively in terms of amplitude of incident wave.

Now ψψ* represents the probability density. Let υ be the velocity of a stream of particles then ψψ*υ represents the current density i.e., number of particles crossing unit area placed perpendicular to the direction of motion. If N_i be the incident flux, then we have by definition

$$N_{i=} |A\,\Psi_i\Psi_i^*|^2\nu = |A\,|^2\nu = |A\,|^2\frac{k_1\hbar}{m} \ldots (11)$$

Similarly the reflected flux N_r is given by the relation

$$N_r = |B|^2\nu = \left|\left(\frac{k_1 - k_2}{k_1 + k_2}\right)\right|^2 |A\,|^2\frac{k_2\hbar}{m} \ldots (12)$$

Now $N_t = \frac{4k_1}{(k_1+k_2)^2}\frac{k_1k_2\hbar}{m}|A|^2 = \frac{4k_1k_2}{(k_1+k_2)^2}N_i$ ... (13)

We have from equation (11), (12) and (13)

$N_t + N_r = N_i$

So quantum mechanically, there is some probability of reflection and some probability of transmission. The reflection coefficient R and transmission coefficient T are defined as under

Reflection coefficient $R = \frac{Magnitude\ of\ reflected\ flux}{Magnitude\ of\ incident\ flux} = \frac{\text{Nr}}{\text{Ni}}$

$$= \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2$$

Transmission coefficient $\text{T} = \frac{Magnitude\ of\ transmitted\ flux}{Magnitude\ of\ incident\ flux} = \frac{\text{Nt}}{\text{Ni}}$

$$= \frac{4k_1k_2}{(k_1 + k_2)^2}$$

So we have $T + R = \frac{4k_1k_2}{(k_1+k_2)^2} + \frac{(k_1-k_2)^2}{(k_1+k_2)^2} = 1$

1.13 APPLICATION OF BARRIER PENETRATION – ALFA DECAY, FIELD IONIZATION AND SCANNING TUNNELLING MICROSCOPE

According to classical ideas, a particle striking a hard wall has no chance of leaking through it. But the behaviour of a quantum particle s different owing to the wave nature associated with it. We know that when an electromagnetic wave strikes at the interface of two media, it is partly reflected and partly transmitted through the interface and enters the second medium. In a similar way the de Broglie wave also has a possibility of getting partly reflected from the boundary of the potential well and partly penetrating through the barrier.

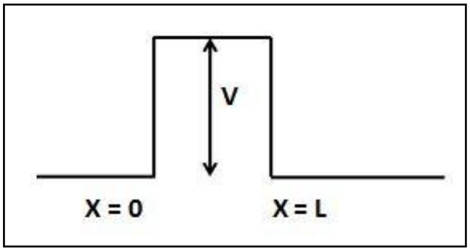


Fig. 1.11

Fig. 1.11 represents a potential barrier of height V and thickness L. A potential barrier is the opposite of a potential well; it is a potential-energy function with a maximum. The potential energy is zero for $x < 0$ and $x > L$ has a value V for $0 < x < L$. An electron of total energy E approaches the barrier from the left. From the view-point of classical physics, the electron would be reflected from the barrier because its energy E is less than V. For the particle to overcome the potential barrier, it must have an energy equal to or greater than V. Quantum mechanics leads to an entirely new result. It shows that there is a finite chance for the electron to leak to the other side of the barrier. We say that the electron tunnelled through the potential barrier and hence in quantum mechanics, the phenomenon is called tunnelling.

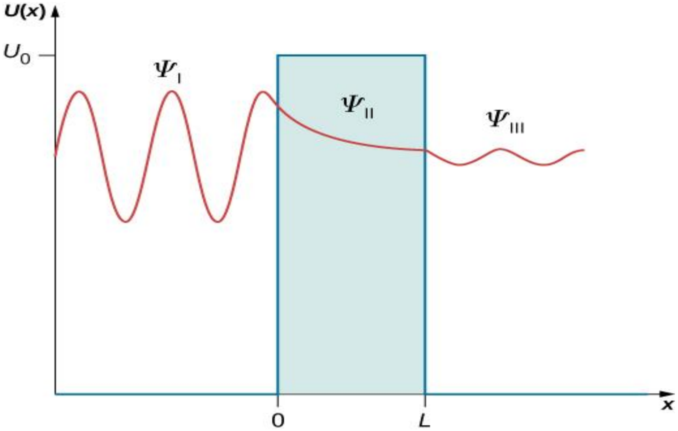


Fig. 1.12

The region around the barrier can be divided into three regions as shown in Fig.(2). We write down the Schrodinger wave equation for the electron wave in the three regions and solve them. The Schrodinger wave equation for regions (I) and (III) has the from

$$\frac{d^2\Psi}{dx^2} + \frac{8\pi^2 mE}{h^2}\Psi = 0 \qquad \text{... (1)}$$

and

$$\frac{d^2\Psi}{dx^2} + \frac{8\pi^2 m}{h^2}(E - V)\Psi = 0 \qquad \text{... (2)}$$

for region (II). The solution of equation (1) is found to be sinusoidal. The solution of equation (2) is found to be exponential. The form of the wave function in region (I), (II) and (III) is also shown in Fig.(2). The wave function Ψ(I) corresponds to the free electron with momentum p = $\sqrt{2mE}$. The wave function Ψ(II) is not zero inside the barrier , but decreases exponentially. Since Ψ(III) is not equal to zero at x = L, there is a finite probability finding the electron in the region III. That means the electron that is initially to the left of the barrier has some probability of being found to the right of the barrier. The wave function Ψ(III) represents the wave transmitted through the barrier and the free electron on the right side of the barrier. The free electron has the same momentum as the incident electron but has smaller amplitude. The electron propagates to the right as a free particle. Thus, it is possible for a particle to penetrate through the potential barrier even if its kinetic energy is less than the height of the potential barrier. The probability that the particle gets through the barrier is called the transmission coefficient.

Quantum mechanics shows that the transmission coefficient is given approximately be

$$T = T_0 e^{-2kl} = T_0 e^{-2l\frac{\sqrt{8\pi^2 m(V-E)}}{h}} \qquad \text{... (3)}$$

where T_0 is a constant close to unity. Equation (3) shows that the probability of particle penetration through a potential barrier depends on the height and width of the barrier.

Tunnelling is significant in many areas of physics. The tunnel diode is a semiconductor diode in which electrons tunnel through a potential barrier. The current can be switched on and off very quickly be varying the height of the barrier, which is done with an applied voltage. The scanning tunnelling electron microscope uses electron tunnelling to produce images of surfaces down to the scale of individual atoms. The quantum possibility of penetrating potential barriers is the basis of the explanation of the α – decay of radioactive nuclei.

1.13.1 Alpha Decay Radioactive

Protons and neutrons in a nucleus have kinetic energy, but it is about 8 MeV less than that needed to get out from attractive nuclear potential (Figure 1). Hence, they are bound by an average of 8 MeV per nucleon. The slope of the hill outside the bowl is analogous to the repulsive Coulomb potential for a nucleus, such as for anα particle outside a positive nucleus. In α decay, two protons and two neutrons spontaneously break away as a ^{4}He unit. Yet the protons and neutrons do not have enough kinetic energy to classically get over the rim.

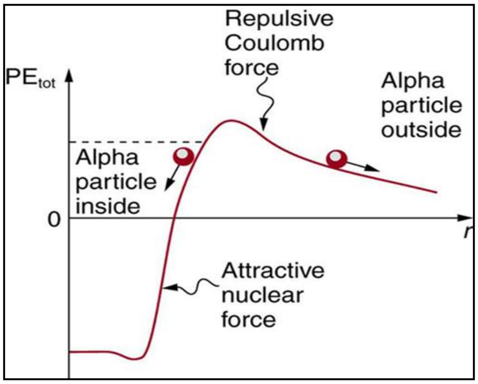


Fig. 1.13

(Nucleons within an atomic nucleus are bound or trapped by the attractive nuclear force, as shown in this simplified potential energy curve. An particle outside the range of the nuclear force feels the repulsive Coulomb force. The α particle inside the nucleus does not have enough kinetic energy to get over the rim, yet it does manage to get out by quantum mechanical tunnelling).

The α particle tunnels through a region of space it is forbidden to be in, and it comes out of the side of the nucleus. Like an electron making a transition between orbits around an atom, it travels from one point to another without ever having been in between (Figure 2). The wave function of a quantum mechanical particle varies smoothly, going from within an atomic nucleus (on one side of a potential energy barrier) to outside the nucleus (on the other side of the potential energy barrier). Inside the barrier, the wave function does not become zero but decreases exponentially, and we do not observe the particle inside the barrier. The probability of finding a particle is related to the square of its wave function, and so there is a small probability of finding the particle outside the barrier, which implies that the particle can tunnel through the barrier.

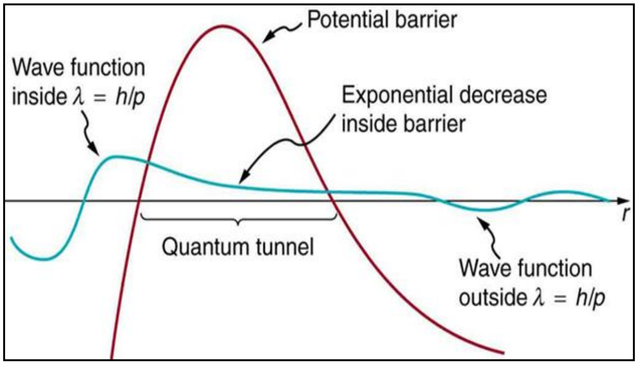


Fig. 1.14

(The wave function representing a quantum mechanical particle must vary smoothly, going from within the nucleus (to the left of the barrier) to outside the nucleus (to the right of the barrier). Inside the barrier, the wave function does not abruptly become zero; rather, it decreases exponentially. Outside the barrier, the wave function is small but finite, and there it smoothly becomes sinusoidal. Owing to the fact that there is a small probability of finding the particle outside the barrier, the particle can tunnel through the barrier)

1.13.2 Field Emission

Field emission is a process of emitting electrons from conducting surfaces due to a strong external electric field that is applied in the direction normal to the surface (Figure 1). As we know from our study of electric fields in earlier chapters, an applied external electric field causes the electrons in a conductor to move to its surface and stay there as long as the present external field is not excessively strong. In this situation, we have a constant electric potential throughout the inside of the conductor, including its surface. In the language of potential energy, we say that an electron inside the conductor has a constant potential energy $U(x) - -U_0$.

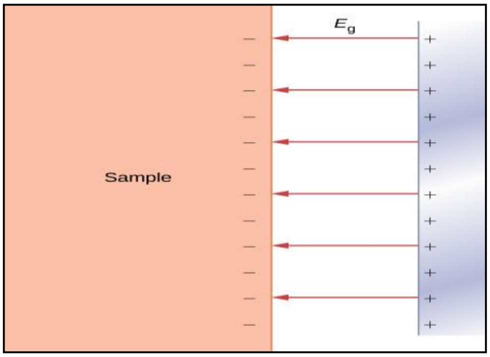


Fig. 1.15

(A normal-direction external electric field at the surface of a conductor: In a strong field, the electrons on a conducting surface may get detached from it and accelerate against the external electric field away from the surface.)

In the situation represented in Figure 1.15, where the external electric field is uniform and has magnitude Eg, if an electron happens to be outside the conductor at a distance x away from its surface, its potential energy would have to be $U(x) = -e\,E_g x$(here, x denotes distance to the surface). Taking the origin at the surface, so that x=0x=0 is the location of the surface, we can represent the potential energy of conduction electrons in a metal as the potential energy barrier shown in Figure 2. In the absence of the external field, the potential energy becomes a step barrier defined by $U(x \leq 0) = -U_0$and by $U(x > 0) = 0$.

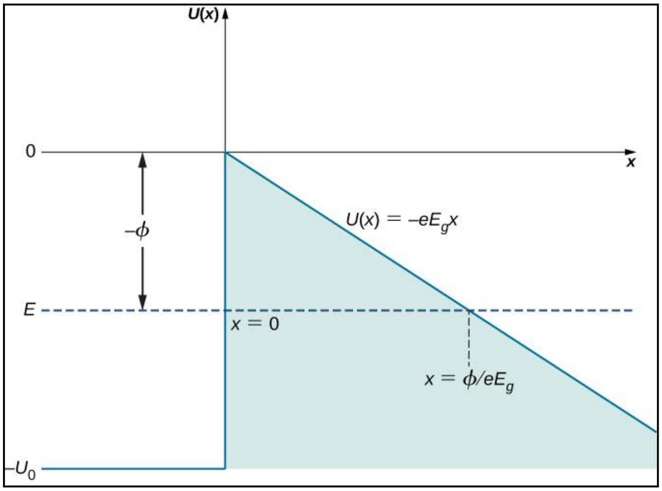


Fig. 1.16

(The potential energy barrier at the surface of a metallic conductor in the presence of an external uniform electric field Eg normal to the surface: It becomes a step-function barrier when the external field is removed. The work function of the metal is indicated by ϕ)

When an external electric field is strong, conduction electrons at the surface may get detached from it and accelerate along electric field lines in a direction antiparallel to the external field, away from the surface. In short, conduction electrons may escape from the surface. The field emission can be understood as the quantum tunnelling of conduction electrons through the potential barrier at the conductor's surface. The physical principle at work here is very similar to the mechanism of α-emission from a radioactive nucleus.

Suppose a conduction electron has a kinetic energy E (the average kinetic energy of an electron in a metal is the work function ϕ for the metal and can be measured, as discussed for the photoelectric effect in , and an external electric field can be locally approximated by a uniform electric field of strength E_g. The width L of the potential barrier that the electron must cross is the distance from the conductor's surface to the point outside the surface where its kinetic energy matches the value of its potential energy in the external field. In Figure 2, this distance is measured along the dashed horizontal line $U(x) = E$ from x=0 to the intercept with $U(x) = -e\, E_g x$, so the barrier width is

$$L = \frac{e^{-1}E}{E_g} = \frac{e^{-1}\emptyset}{E_g}$$

We see that L is inversely proportional to the strength E_g of an external field. When we increase the strength of the external field, the potential barrier outside the conductor becomes steeper and its width decreases for an electron with a given kinetic energy. In turn, the probability that an electron will tunnel across the barrier (conductor surface) becomes exponentially larger. The electrons that emerge on the other side of this barrier form a current (tunnelling-electron current) that can be detected above the surface. The tunnelling-electron current is proportional to the tunnelling probability. The tunnelling probability depends nonlinearly on the barrier width L, and L can be changed by adjusting E_g. Therefore, the tunnelling-electron current can be tuned by adjusting the strength of an external electric field at the surface. When the strength of an external electric field is constant, the tunnelling-electron current has different values at different elevations L above the surface.

1.13.3 Scanning Tunneling Microscopy (STM)

A metal tip made up of tungsten is placed between a very small distances above a conducting or semiconducting surface. This distance acts as a potential barrier for tunnelling. The space between the tip and the surface normally is vacuum. When electrons tunnel from the metal tip to the surface, a current is created and monitored by a computer (Figure 1). The current depends on the distance between the tip and the surface, which is controlled by a piezoelectric cylinder. When there is a strong current, the tip will move away from the surface. The increase of the potential barrier will decrease the probability of tunnelling and decrease the current. If the current becomes too weak, the tip moves closer to the surface. The potential barrier will be reduced and the current will increase. The variations in the current as the tip moves over the sample are reconstructed by the computer to produce topological image of the scanned surface.

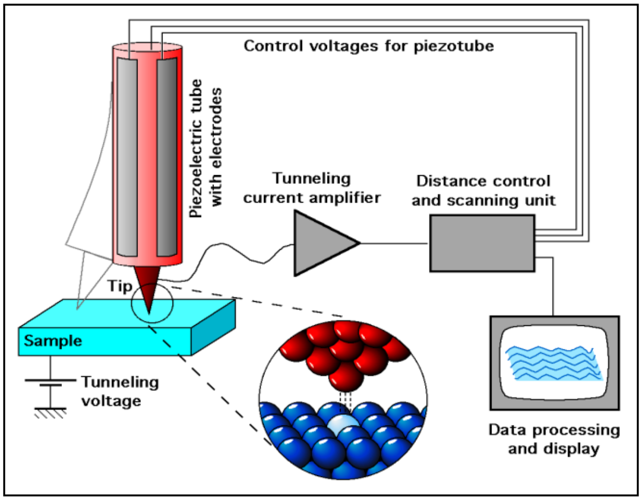


Fig. 1.17 Schematic diagram of a scanning tunnelling microscope

Just to know

Erwin Schrodinger

(1887 – 1961)

Erwin Schrodingerwas a Nobel Prize-winning Austrian physicist who developed a number of fundamental results in the field of quantum theory: the Schrödinger equation provides a way to calculate the wave function of a system and how it changes dynamically in time.

In addition, he was the author of many works in various fields of physics: statistical mechanics and thermodynamics, physics of dielectrics, colour theory, electrodynamics, general relativity, and cosmology, and he made several attempts to construct a unified field theory. In his book *What Is Life?* Schrödinger addressed the problems of genetics, looking at the phenomenon of life from the point of view of physics. He paid great attention to the philosophical aspects of science, ancient and oriental philosophical concepts, ethics, and religion. He also wrote on philosophy and theoretical biology. He is also known for his "Schrödinger's cat" thought-experiment.

Basic Quantum Physics

Problems

1. Calculate the minimum energy of an electron can possess in an infinitely deep potential will of width 4nm.

   Given:

   a = 4nm; n = 1 (for minimum energy)

Solution:

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

$$E = \frac{1^2(6.625 \times 10^{-34})^2}{8 \times 9.11 \times 10^{-31}(4 \times 10^{-9})^2}$$

$$E = 3.7639 \times 10^{-21}\, J \; or \; E = 0.0235\, eV$$

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

   Given:

   a = 0.1 nm; n = 1 (for ground energy)

Solution:

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

$$E = \frac{1^2(6.625 \times 10^{-34})^2}{8 \times 9.11 \times 10^{-31}(0.1 \times 10^{-9})^2}$$

$$E = 6.0223 \times 10^{-18}\, J$$

For sixth state, n = 6

$$E = \frac{6^2(6.625 \times 10^{-34})^2}{8 \times 9.11 \times 10^{-31}(0.1 \times 10^{-9})^2}$$

$$E = 2.1680 \times 10^{-16}\, J$$

The energy required to excite the electron from its ground state to the sixth excited state is $E = E_6 - E_1$

$$E = 2.1680 \times 10^{-16} - 6.0223 \times 10^{-18}$$

$$E = 2.1078 \times 10^{-16}\, J$$

$$E = 1317\, eV$$

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

Solution:

$$\Delta p_x \, \Delta x \approx h$$

$$\Delta p_x \approx \frac{h}{\Delta x} \approx \frac{6.6 \times 10^{-34}}{4 \times 10^{-10}}$$

$$= 1.65 \times 10^{-24}\, kg \frac{m}{sec}$$

Basic Quantum Physics

QUESTIONS

Part A:

1. Define electromagnetic wave
2. Define the following
   a) Radio waves
   b) Microwaves
   c) infrared waves
   d) Visible rays
   e) Ultraviolet rays
   f) X- Rays
   g) Gamma rays
3. Explain Photo electric effect.
4. Define Compton scattering
5. Define photon and write any four properties of photon.
6. Write postulates of Bohr Atom model.
7. Define electron diffraction.
8. Write dual nature of the particle.
9. Define De-Broglie wavelength.
10. Deing wave function with the basic equation.
11. What is Heisenberg Uncertainty principle?
12. Write the importance of Heisenberg Uncertainty principle.
13. State Schroedinger's time dependent wave equation.
14. State Schroedinger's time independent wave equation.
15. Write down Schroedinger's time dependent and independent wave equations for matter waves.

16. What is operator?
17. What do you mean by eigen values and eigen functions of the operators?
18. What is transmission and reflection coefficient?
19. What is quantum mechanical tunnelling?
20. Write the principle behind the STM.

Part B:

1. Explain Photo electric effect with its experimental evidences.
2. Explain in detail about Franck-Hertz experiment to give Quantum mechanical evidences for Bohr Atom Model.
3. State Compton Scattering and write relationship between relativistic energy and momentum of the particle.
4. What is uncertainty principle? Explain how it is the out come of the wave description of a particle.
5. Derive Schroedinger's time dependent wave equations for matter waves.
6. Derive stationary state wave equations for matter waves.
7. State Schroedinger's time independent wave equation for a one-dimensional case. Use it to prove that a particle enclosed in a one-dimensional box has quantised energy values.
8. Obtain expression for Eigen function of particle in one dimensional potential well of infinite height.
9. A particle of some mass number and total energy E moves from a region of zero potential to a region of constant potential V_0. Derive an expression for reflection and transmission coefficients when $E<V_0$.
10. Explain in detail potential barrier and quantum mechanical tunnelling.
11. Discuss in detail the applications of quantum tunnelling.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF PHYSICS

UNIT – II - PHYSICS FOR ENGINEERS-SPHA1101

UNIT 2

PHYSICS OF SOLIDS

2.0INTRODUCTION

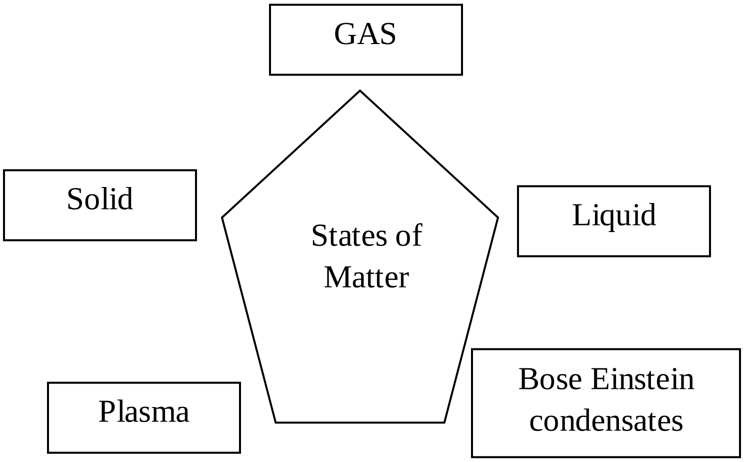


Fig.2.1States of Matter

There are five known phases in states of matter are, Gas, Liquid, Solids, Plasma and Nose Einstein condensates.

Gas has no definite volume and no definite shape. Gas particles have a great deal of space between them. Hence the atoms are more free to move. Gas particles have high kinetic energy.

The increasing pressure reduces the space between the particles and the pressure exerted by their collisions increases. Liquids have an indefinite shape. The liquid particles are not held in a regular arrangement, but are still very close to each other so liquids have a definite volume.

Hence the atoms are free to move than solids. In the liquid phase, the particles of a substance have more kinetic energy than those in a solid. Liquids, cannot be compressed under pressure.

Solids have a definite shape. The atoms are tightly packed in a solid. Hence the atoms are unable to move freely. Particles of a solid have very low kinetic energy. The increasing pressure will not compress the solid to a smaller volume. Based on the arrangement of atoms, the solids are further classified into three major categories are (i) Crystalline solids (ii) Semi crystalline solids and (iii) amorphous solid.

Plasma is a charged gas, (neutral molecules and atoms) with strong Coulomb interactions**.** Plasma is not a common state of matter on Earth, but the most common state of matter in the universe.Plasma consists of highly charged particles with extremely high kinetic energy. The nobel gases such as helium, neon, argon, krypton, xenon and radon are often used to make glowing signs by using electricity to ionize them to the plasma state. Stars are essentially superheated balls of plasma.

In 1995, technology enabled scientists to create a new state of matter, the Bose-Einstein condensate (BEC). Using a combination of lasers and magnets, Eric Cornell and Carl Weiman cooled a sample of **Rubidium** to within a few degrees of absolute zero. At this extremely low temperature molecular motion comes very close to stopping altogether. Since there is almost no kinetic energy being transferred from one atom to another, the atoms begin to clump together. There are no longer thousands of separate atoms, just one “super atom”. BECs are also used to simulate conditions that might apply in black holes.

Let us briefly discuss about solids in our further discussions.

2.1STRUCTURE OF SOLIDS

Solids are classified into following categories based on the arrangement of atoms. There are:

- **Amorphous solids**

An amorphous solid is a substance whose constituents do not possess an orderly arrangement and isotropic in nature. Important examples of amorphous solids are glass and plastics. Although amorphous solids consist of microcrystalline substance but the orderly arrangement is restricted to very short distances.They do not have definite heat of fusion and sharp melting point.

- **Semi (Poly) crystalline solids**

Semi-crystalline materials have a highly ordered molecular structure with sharp melt points. They have excellent chemical resistance.

- **Crystalline solids**

A crystalline solid is a substance whose constituents possess an orderly arrangement in a definite geometric pattern. Some very common examples of crystalline substances are sodium chloride, sugar and diamond. Crystalline solids exhibit definite heats of fusion.

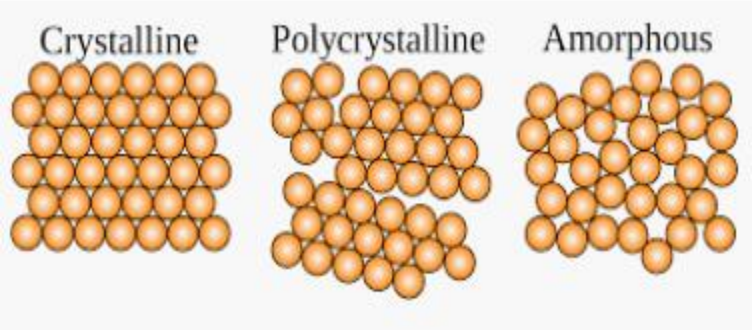
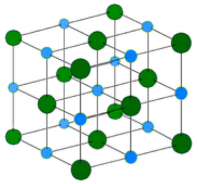
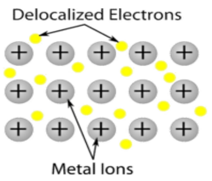
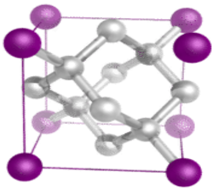
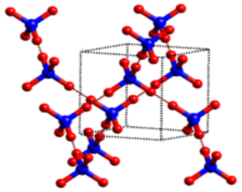


Fig. 2.2Types of solids

Crystalline substances can be described by the types of particles in them and the types of chemical bonding that takes place between the particles. Based on chemical bonding the types of solids as follows:

2.1.1 Types of Crystalline Solids

There are four types of crystals: (1) **ionic** (2) **metallic** (3) **covalent** network, and (4) **molecular**. Properties and several examples of each type are listed in the following table and are described in the table below.

	Ionic Solid	Metallic Solid	Covalent Solid	Molecular Solid
1	The ionic crystal structure consists of alternating positively-charged cations and negatively-charged anions The ions may either be monatomic or polyatomic.	Metallic crystals consist of metal cations surrounded by a "sea" of mobile valence electrons These electrons, also referred to as delocalized electrons, do not belong to any one atom, but are capable of moving through the entire crystal.	A covalent network crystal consists of atoms at the lattice points of the crystal, with each atom being covalently bonded to its nearest neighbor atoms.	Molecular crystals typically consist of molecules at the lattice points of the crystal, held together by relatively weak intermolecular forces. Molecular solids are three types (i) Polar (ii) Non polar (iii)Hydrogen bonded
2	They have high melting points	They have wide range of melting points	They have high melting and boiling points.	They have low melting and boiling points.
3	Hard but brittle; shatter under stress	soft	very hard and brittle	Easily deformed under stress; ductile and malleable
4	Poor conductors of heat and electricity	Good conductors of heat and electricity than ionic crystal	Poor conductors of heat and electricity	Good conductors of heat and electricity
5	Relatively dense	low density	low density	high density
6	 NaCl crystal	 Iron	 Diamond	 Ice crystal structure

Several theories are successfully explained the various properties of solids and it is explained as follows.

2.2 ELECTRON THEORY OF SOLIDS

The electron theory of solids explains the structures and electrical, thermal properties of solids through their electronic structure. This theory also explains the binding in solids, behavior of conductors, semiconductors and insulators, Ferromagnetism, electrical and thermal conductivities of solids, elasticity, cohesive and repulsive forces in solids etc., the theory has been developed in three main stages, are.

(I) Classical free electron theory of metals

(II) Quantum free electron theory of metals

(III) Band theory of metals (or) Brillouin zone theory

The brief discussion about these theories as follows;

2.2.1 Classical free electron theory of metals

The classical free electron theory was introduced by P. Drude in 1900 and developed by Lorentz in 1909 to explain electrical conduction in metals. This theory has some assumptions / Postulates; they are:

1. The valence electrons of metallic atoms are free to move in the spaces between ions from one place to another place within the metallic specimen similar to gaseous molecules so that these electrons are called free electron gas. These electrons also participate in electrical conduction; hence they are called conduction electrons.
2. There is no interaction between these conduction electrons.
3. The interaction of free electrons with ion cores is negligible.
4. The free electrons find uniform electric field of positive ions and that of electrons in the metal.

5. Since the electrons are assumed to be a perfect gas they obey classical kinetic theory of gasses.
6. Classical free electrons in the metal obey Maxwell-Boltzmann statistics.

2.2.2 Postulates of free electron theory in absence of electric field:

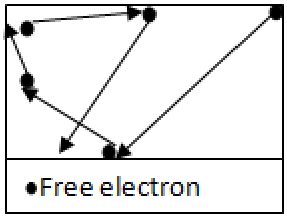


Fig. 2aFree motion of electron

The free electrons are moving randomly shown in Fig.2a within the boundaries of a metal similar to the gas molecules. The total energy of the electron is assumed to be purely about its kinetic energy because of its zero potential energy

Due to the free motion of free electrons the electrons are collides with each other elastically without losing their energy, called as free electrons.

2.2.3 In presence of electrical field:

To study electrical conductivity

In the absence of an external electric field, the free electrons move in random directions, similar to gaseous molecules in a container so that there is no net resultant motion of electrons along any direction.

When an electric field is applied to the conductor, the electrons are subjected to acceleration. The drift velocity of the electrons will depend upon the applied field, greater is the drift velocity and consequently, higher would be the number of electrons per second through unit area, i.e., the current passing through the conductor will be high.

To obtain an expression for electrical conductivity, consider a metallic rod of length *l* and area of cross section '*A*' as shown in Fig. 2b.

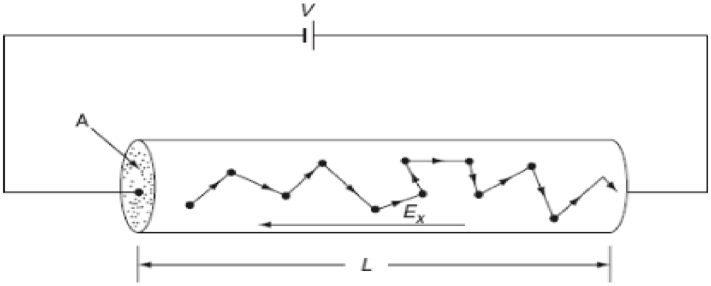


Fig. 2b Electron drifting in electric field (E_x)

The root mean square velocity of electrons is obtained by applying kinetic theory of gases and equipartition law of energy. The pressure '*P*' of the gas is:

$$P = \frac{1}{3}\rho\bar{C}^2 = \frac{1}{3}mn\bar{C}^2 \qquad (1)$$

where *m* is mass of an electron, '*n*' is the number of free electrons per unit volume, '*ρ*' is the density of an electron gas and $\bar{C}$ is root mean square velocity.

For molar volume (V_m) of the metal,

$$P = \frac{1}{3}\rho\bar{C}^2 = \frac{1}{3}m\frac{N_A}{V_m}\bar{C}^2 \qquad (2)$$

where N_A is Avogadro's number

From Equation (2)

$$PV_m = \frac{1}{3}mN_A\bar{C}^2 = R_uT \ \text{(or)} \quad mn\bar{C}^2 = \frac{3R_uT}{N_A} = 3K_BT \qquad (3)$$

where R_u is universal gas constant, *T* is absolute temperature and K_B is Boltzmann's constant.

From Equation (3)

$$\bar{C}^2 = \sqrt{\frac{3K_BT}{m}} \tag{4}$$

Now a potential difference (*V*) is applied across the ends of the rod, kept parallel to X-direction, then an electric field, E_x acts on the electrons, then the electrons acquire a constant velocity v_x in the opposite direction to that of electric field. This constant velocity is called drift velocity, this velocity is superimposed on the thermal velocity ($\bar{C}$) of the electrons.

If 'R_x' is the resistance of the rod, then the current I_x through it is:

$$I_x = \frac{V_x}{R_x} = AJ_x \tag{5}$$

where J_x is the current density along X-direction

$$\text{(or)} \quad J_x = \frac{V_x}{AR_x} = \frac{E_x l}{A} \times \frac{A}{\rho_x l} = \frac{E_x}{\rho_x} = \sigma_x E_x \tag{6}$$

where E_x = intensity of electric field; ρ_x = resistivity and σ_x = electrical conductivity of the metal.

The field E_x produces a force − eE_x on each electron, due to this force the acceleration on *i*th electron is:

$$a_{ix} = \frac{dv_{ix}}{dt} = -\frac{eE_x}{m} \tag{7}$$

wherev_{ix} is the drift velocity of *i*th electron along X-direction. Since the electrical force eE_x is the same on all electrons, so the average acceleration is:

$$\frac{d}{dt}\langle v_x \rangle = -\frac{eE_x}{m} \tag{8}$$

where$\langle v_x \rangle$ is the average velocity of *n* free electrons in unit volume of metal, given by:

$$\langle v_x \rangle = \frac{1}{n} \sum_{i=1}^{n} v_{ix} \qquad (9)$$

The current density along X-direction is:

$$J_x = n \langle v_x \rangle (-e) = -ne \langle v_x \rangle \qquad (10)$$

The minus sign shows that J_x is in the opposite direction to that of $\langle v_x \rangle$. The average velocity $\langle v_x \rangle$ can be obtained by considering the acceleration and retardation in steady state.

In addition to acceleration, $[eE_x/m]$ the electrons get retarded due to collisions of electrons with lattice ions. This is referred to as electron lattice scattering.

Where,

$$\langle v_x \rangle = -\left(\frac{eT_x}{m} \right) E_x \qquad (11)$$

where τ_x is known as relaxation time. It is defined as the time taken by an electron to decay its drift velocity to 1/*e* of its initial value.

Equation (11) shows that the drift velocity acquired by the electron is proportional to applied electric field and the proportionality constant $(e\tau_x/m)$ is known as their mobility (μ_x). It is defined as the drift velocity produced per unit applied electric field and it is given by:

$$\mu_x = \frac{\langle v_x \rangle}{E_x} = \frac{eT_x}{m} \qquad (12)$$

Finally, comparing Equations $J_x = \sigma_x E_x$ and $J_x = -ne \langle v_x \rangle$, we have:

$$\sigma_x E_x = -ne \langle v_x \rangle \qquad (13)$$

Substituting Equation (11) in (13)

$$\sigma_x E_x = -ne\left(-\frac{eT_x}{m}\right)E_x$$

$$\therefore \sigma_x = \frac{ne^2}{m}\tau_x \qquad (14)$$

Substituting Equation (12) in (14) gives:

$$\sigma_x = ne\mu_x \qquad (15)$$

(or) $$\text{resistivity, } \rho_x = \frac{1}{\sigma_x} = \frac{1}{ne\mu_x} \qquad (16)$$

Equations (14), (15) and (16) represent electrical conductivity and resistivity of a material.

Basic definitions:

- Drift velocity (v_d): It is defined as the average velocity acquired by the free electron in a particular direction due to the application of electric field.
- Relaxation time (τ): It is the time taken by the free electron to reach its equilibrium position from its disturbed position in the presence of applied field.
- Collision time (τ_c): It is the average time taken by a free electron between two successive collisions.
- Mean free path (λ): The average distance travelled between two successive collisions is called mean free path

Physics of Solids

Thermal Conductivity (K)

Definition:

Thermal conductivity of the material is defined as the amount of heat conducted per unit area per unit time maintained at unit temperature gradient.

$$Q = -K\frac{dT}{dx}$$

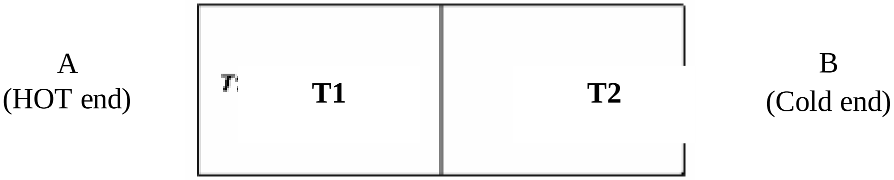
where, dT/dx = temperature gradient

Q = heat flux or heat density

K = the coefficient of thermal conductivity of the material.

Expression for Thermal Conductivity of a metal:

On the basis of kinetic theory, if two temperatures are equal, if T1 = T2 , then there is no transfer of energy. If T1 >T2 , there is transfer of energy from A to B (Fig.)



Consider a uniform rod AB with the temperature T_1 at end A and T_2 at the other end.

The two cross sections A and B are separated by a distance λ. Heat flows from hot end A to the cold end b by the electrons. The amount of heat (Q) conducted by the rod from the end A to be of length λ is given by,

$$Q = \frac{KA(T_1 - T_2)t}{\lambda} \qquad (17)$$

∴The coefficient of conductivity per unit area per unit time

$$K = \frac{Q}{(T_1 - T_2)}\lambda \tag{18}$$

where, K = coefficient of thermal conductivity

A = Area of cross section of the rod In Collision, the electrons near A loses their kinetic energy while the electrons near B gains the energy.

The average kinetic energy of the electron near A

$$A = \frac{3}{2}K_BT \tag{19}$$

The average kinetic energy of the electron near B

$$B = \frac{3}{2}K_BT_2 \tag{20}$$

Where, K_B = Boltzmann constant

∴Excess kinetic energy carried by the electron from A to B (eqn. 19 - eqn.20)

$$= \frac{3}{2}K_BT_1 - \frac{3}{2}K_BT_2$$

$$= \frac{3}{2}K_B(T_1 - T_2) \tag{21}$$

Let us assume that there is equal probability for the electrons to move in all the 6 directions as shown in below Fig.2c. Since each electron travels with thermal velocity ‘v’ , if n is the free electron density, then on an average 1/6 nv.

Electrons will travel in any one direction. Then, The number of electrons crossing unit area per second in one direction from A to B

$$=\frac{1}{6}\,nv \qquad (22)$$

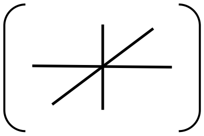


Fig.2c (6 directions)

Excess of energy transferred from A to B for unit area in unit time

$$=\frac{1}{6}\,nv[\text{K. Eat A} - \text{K.Eat B}] \qquad (23)$$

We know that Kinetic energy K.E

Therefore the total energy transferred

$$=\frac{1}{2}\,mv^2 = \frac{3}{2}K_BT$$

$$=\frac{1}{6}\,nv\left[\frac{3}{2}K_BT_1 - \frac{3}{2}K_BT_2\right]$$

$$=\frac{1}{6}\,nv\ \frac{3}{2}K_B[T_1 - T_2]$$

$$=\frac{nvK_B(T_1 - T_2)}{4} \qquad (24)$$

Similarly, deficiency of energy carried from B to A for unit area in unit time

$$=\frac{nvK_B(T_1 - T_2)}{4} \qquad (25)$$

Hence net energy transferred from A to B per unit area per unit time (or) heat flux Q is given by If K be the thermal conductivity of the material, then the transfer of energy per unit area per unit time is given by

$$Q = \frac{nvK_B(T_1 - T_2)}{4} - \left(- \frac{nvK_B(T_1 - T_2)}{4} \right)$$

$$Q = \frac{nvK_B(T_1 - T_2)}{4} + - \frac{nvK_B(T_1 - T_2)}{4}$$

i.e. $$Q = \frac{nvK_B(T_1 - T_2)}{4} \qquad (26)$$

From Eqns.(26) and (27)

$$-\frac{K(T_1 - T_2)}{\lambda} = \frac{nvK_B(T_1 - T_2)}{2} \qquad (27)$$

$$\therefore K = \frac{nvK_B\lambda}{2} \qquad (28)$$

Where λ is the mean free path of an electron.

We know that for metals, Relaxation time = Collision time

$$\text{i.e., } \tau = \tau_c = \frac{\lambda}{v}$$

(or) $$\tau v = \lambda \qquad (29)$$

Substituting eqn.(29) in equation (28), we get

$$K = \frac{nvK_B\tau}{2} \qquad (30)$$

Equation (30) is the classical expression for thermal conductivity.

Wiedemann – Franz law (Derivation) Statement

The ratio between the thermal conductivity and the electrical conductivity of a metal is directly proportional to the absolute temperature of the metal. This ratio is constant for all metals.

i.e., $\frac{K}{\sigma} \infty T$ (or) $\frac{K}{\sigma} = LT$ (A)

Where L is known as Lorentz number whose value is, 2.44×10^{-8} $W\Omega K^{-2}$ at 293 K

Derivation

Wiedemann-Franz law can be derived using the expressions of thermal conductivity and electrical conductivity of metal. Electrical conductivity of metals is given by,

$$\sigma = \frac{ne^2\tau}{m} \qquad (31)$$

From eqn (14)

Thermal conductivity is given by,

$$K = \frac{nv^2K_B\tau}{4} \qquad (32)$$

From eqn (30)

From eqn (31) and eqn (32) and (A)

$$\frac{\text{Thermal conductivity}}{\text{Electrical conductivity}} = \frac{K}{\sigma} = \frac{\frac{1}{2}nv^2K_B\tau}{\frac{ne^2\tau}{m}}$$

$$\frac{K}{\sigma} = \frac{1}{2}\frac{mv^2K_B}{e^2} \qquad (33)$$

We know that the kinetic energy of an electron

$$\frac{1}{2}mv^2 = \frac{3}{2}K_BT \tag{34}$$

From eqn (33) and eqn (34) we get,

$$\frac{K}{\sigma} = \frac{3}{2}\frac{K_BTK_B}{e^2}$$

$$\frac{K}{\sigma} = \frac{3}{2}\left(\frac{K_B}{e}\right)^2 T \tag{35}$$

(or) $$\frac{K}{\sigma} = LT \tag{36}$$

Where , L is a constant known as Lorentz Number. By comparing (35) and (36);

$$L = \frac{3}{2}\left(\frac{K_B}{e}\right)^2$$

(or) $$\frac{K}{\sigma} \infty T \tag{37}$$

Thus the ratio of the thermal conductivity to electrical conductivity of a metal is directly proportional to the absolute temperature of the metal.

This proves Wiedemann-Franz law for a conducting material.

Lorentz Number: Lorentz number is a constant given by

$$L = \frac{3}{2}\left(\frac{K_B}{e^2}\right)^2$$

$K_B = 1.39 \times 10^{-23}$ JK^{-1}

$e = 1.6021 \times 10^{-19}$ Coulomb

$$\therefore L = \frac{3}{2}\left(\frac{1.38 \times 10^{-23}}{1.6021 \times 10^{-19}}\right) = 1.12 \times 10^{-8}\, W\Omega K^{-2}$$

i.e. $L = 1.12 \times 10^{-8}\ \mathrm{W\Omega K^{-2}}$

where, It is found that the classical value of Lorentz number is only half of the experimental value i.e., $2.44 \times 10^{-8}\ \mathrm{W\Omega K^{-2}}$

This discrepancy in the experimental and theoretical value of Lorentz number is the failure of classical theory. This discrepancy is rectified by quantum theory.

Merits of classical free electron theory:

- Classical theory successfully explains electrical and thermal conductivities of metals, the increase of resistance with temperature, opacity, luster, etc.,
- It verifies Ohm's law
- It is used to derive Wiedemann Franz law
- It is also used to explain optical properties

Drawbacks ofclassical free electron theory:

- It is a macroscopic theory
- There is an controversy in Classical theory and quantum theory; Classical theory states that all free electrons will absorb energy but quantum theory states that only few electron will absorb energy.
- This theory cannot explain the Compton effect , photo electric effect and about magnetism.
- The theoretical and experimental electronic specific heat value is not agreed well.
- According to classical theory, the electrical conductivity (σ) of a metal is inversely proportional to its square root of its absolute temperature (T). i.e.,

$\sigma \propto \frac{1}{\sqrt{T}}$. But practically, it has been observed that the electrical conductivity is inversely proportional to its absolute temperature. i.e., $\sigma \propto \frac{1}{\sqrt{T}}$.

- The Lorentz number is mismatched in classical theory and it is rectified in quantum theory.

2.2.4 Quantum free electron theory

- This theory is proposed by Sommerfeld in 1928 and It is a Microscopic theory.
- It is explained that the electrons move in a constant potential and it obeys quantum laws
- In this theory it is assumed that the electron is moving with a constant potential
- The mass of an electron is constant, when it moves in constant potential
- This theory states that the electrons are moving with higher velocities the potential energy of an electron becomes zero.
- This theory successfully explains electrical conductivity, specific heat, thermionic emission and paramagnetism.
- The quantum theory verified Wiedemann Franz law and proved the value of Lorentz number by using the concept of effective mass of an electron.
- This theory failed to explain some of the physical properties are (i) it fails to explain the difference between conductors, insulators and semiconductors, (ii) positive Hall coefficient of metals and (iii) lower conductivities of divalent metals than monovalent metals.

Wiedemann Franz law by quantum theory:

In quantum theory the mass of an electron is replaced by effective mass of an electron m*,

Hence the expression of electrical conductivity eqn. (14) becomes,

$$\sigma = \frac{ne^2\tau}{m^*} \qquad (1)$$

The thermal conductivity expression eqn (30) becomes,

$K = 2\ (\pi/3)^2\ \times 3/2\ [nK_B{}^2T\tau/m]$

$[1/2mv^2 = 3/2\ K_BT;$ Hence $v^2 = 3\ K_BT/m]$

$K = \pi^2 nK_B{}^2T\tau/3m^* \qquad (2)$

By using (A), (1) and (2)

$K/\sigma = \pi^2/3\ (K_B/e)^2T \qquad (3)$

Hence $L= \pi^2/3\ (K_B/e)^2= 2.44 \times 10^{-8}\ W\Omega K^{-2}$

This shows that, quantum theory verified Wiedemann Franz law and good agreement with Lorentz number.

2.2.5Brillouin Zone theory / Band Theory

- This theory is proposed by Bloch in 1928.
- It is explained the concept of electron moves in a periodic potential.
- According to the zone theory, the mass of an electron varies when it moves through the periodic potential and it is called as effective mass of an electron.

This theory also explains the mechanism of Conductors, semiconductors and insulators based on bands and hence called as band theory.

Let us discuss about formation of bands in solids as follows.

2.3ORIGIN OF ENERGY BANDS

Generally solid consists of more number of atoms and each atom consists of the central nucleus of positive charge around which small negatively charged particles called electrons revolve in different paths or orbits. An Electrostatic force of attraction between electrons and the nucleus holds up electrons in different orbits.

An isolated atom possesses discrete energies of different electrons. If two isolated atoms are brought to very close proximity, then the electrons in the orbits of two atoms interact with each other. So, that in the combined system, the energies of electrons will not be in the same level but changes and the energies will be slightly lower and larger than the original value. So, at the place of each energy level, a closely spaced two energy levels exists. If 'N' number of atoms are brought together to form a solid and if these atoms' electrons interact and give 'N' number of closely spaced energy levels in the place of discrete energy levels, it is known as bands of **allowed energies/band.** Between the bands of allowed energies, a small region does not allow any energy levels. Such regions are called **forbidden energy band/ energy gap** (E_g). The allowed bands of energies and forbidden bands of energies are present alternatively one after another for the electrons of a solid (Fig.2.3).

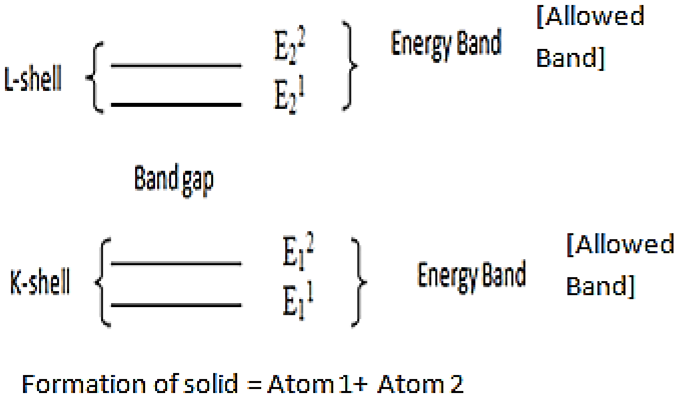


Fig. 2.3Band formation in solids

When two atoms of equal energy levels are brought closer together, the k shell energy levels of individual atoms lie in the energy level E_1 and splits into E_1^1 and E_1^2 become as **energy band.** Similarly for L-shell energy levels lies in the energy level E_2 and splits into E_2^1 and E_2^2. Similarly when three atoms are brought together, the energy levels are in K – shell and L-shell are E_1^1 , E_1^2 and E_2^1, E_2^2 respectively.

Similarly for K, L,M and N shells are E_1^1 , E_1^2,E_1^3 , E_1^4 and E_2^1, E_2^2, E_2^3, E_2^4 respectively. These type of transformations from the original energy levels into two (or) more energy levels is known as energy level splitting.

In example Lithium atom consists of 3 electrons. Among them the two electrons are in the same energy state (1s) and the other one is in the higher energy state (2s). In a solid of lithium contains N number of atoms, the lower energy level forms a band of 2N electrons occupying N different energy levels are shown in Fig.2.4. The higher energy level can form a band of 2N electrons (completely filled). But only N numbers of electrons are present in the next energy level. Therefore the band is half-filled band. The unoccupied band of the solid corresponds to the un occupied excitation level of the isolated atom.

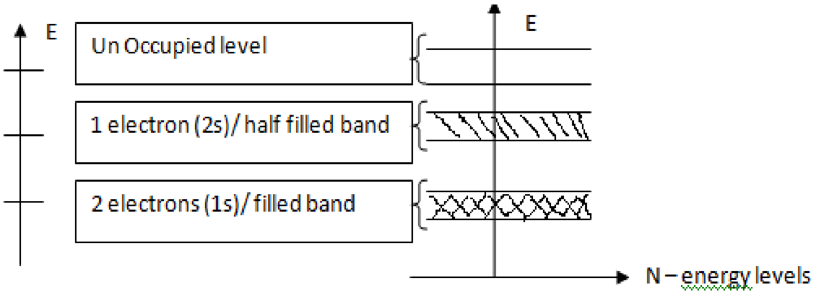


Fig.2.4 (a) Isolated atom and (b) Solid of N atoms

The width of the energy band depends on the relative freedom of electrons in the crystal. The electrons in lower filled energy (closely bounded with nucleus) band do not have any freedom to move in the crystal. The electrons in half filled energy level (loosely bound with nucleus) are free to move inside the crystal. Generally the half-filled band is called **valence band** and the upper unoccupied (unfilled) band is called **conduction band**. In between the valence band and conduction the gap is known as energy gap

The differences between metals, semiconductors and insulators can be made based on (i) the width of the energy gap (ii) number of effective electrons (The availability of number of free electrons that participate in electrical conduction per unit volume when an electric field is applied. These free electrons are called **effective electrons)**.

2.4 BAND STRUCTURE OF CONDUCTOR, SEMICONDUCTORS, INSULATORS, HALF METALS AND SEMI METALS

2.4.1 Conductors:

In case of conductors, there is no forbidden band and the valence band and conduction band overlap each other (Fig.2.5). Low resistive materials and it is about $10^{-8}\Omega m$ are generally called as conducting materials. These materials have high electrical and thermal conductivity. The conducting property of a solid is not a function of total number of the electrons in a metal and it is about the number of free/valence electrons called as conduction electrons. Hence in metals the electrical conductivity depends on the number of free electrons. Here, plenty of free electrons are available for electric conduction. The electrons from valence band freely enter in the conduction band due to overlapping of bands.

Examples: Iron, Copper etc.,

2.4.2 Semiconductors:

In semiconductors, the forbidden band is very small (Fig. 2.5). Germanium and silicon are the best examples of semiconductors. In Germanium the forbidden band is of the order of 0.7eV while in case of silicon, the forbidden gap is of the order of 1.1eV. Actually, a semi-conductor material is one whose electrical properties lie between insulators and good conductors. At 0°K, there are no electrons in conduction band and the valence band is completely filled. When a small amount of energy is supplied, the electrons can easily jump from valence band to conduction band due to minimum energy gap. For example, when the temperature is increased, the forbidden band is decreased so that some electrons are liberated into the conduction band. In semi-conductors, the conductivities are of the order of 10^2 ohm-meter.

Examples: Silicon, Germanium etc.,

The detailed description about the types of semiconductor as follows in section 2.7.

2.4.3 Insulator

In case of insulators, the forbidden energy band is wide (Fig.2.5). Due to this fact, electrons cannot jump from valence band to conduction band. In insulators, the valence electrons are bound very tightly to their parent atoms. For example, in case of materials like glass, the valence band is completely filled at 0K and the energy gap between valence band and conduction band is of the order of 6 eV. Even in the presence of high electric field, the electrons do not move from valence band to conduction band. When a very large energy is supplied, an electron may be able to jump across the forbidden gap. Increase in temperature enables some electrons to go to the conduction band. This explains why certain materials which are insulator become conductors at high temperature. The resistively of insulators is of the order of 10^7 ohm-meter.

Examples: Wood, Glass etc.,

At absolute zero (0K) the energy levels are completely occupied by electrons up to a certain level. Above that the energy levels are completely unoccupied by electrons. This highest level which is completely filled by electrons is called **Fermi level**. In conductor the Fermi level lies in the permitted band (since the valence and conduction band overlap without an energy gap). In Semiconductor the Fermi level lies in the small energy gap. In Insulator the Fermi level lies exactly in the middle of energy gap.

The **Fermi energy** (E_F) is described as the highest energy that the electrons assumes at a temperature of 0 K (Fig.2.5). The Fermi energy is the difference in energy, mostly kinetic. To put this into perspective one can imagine a cup of coffee and the cup shape is the electron band; as one fills the cup with the liquid the top surface increases. This can be compared to the Fermi energy because as electrons enter the electron band the Fermi energy increases. The Semiconductors act differently than the above stated analogy.

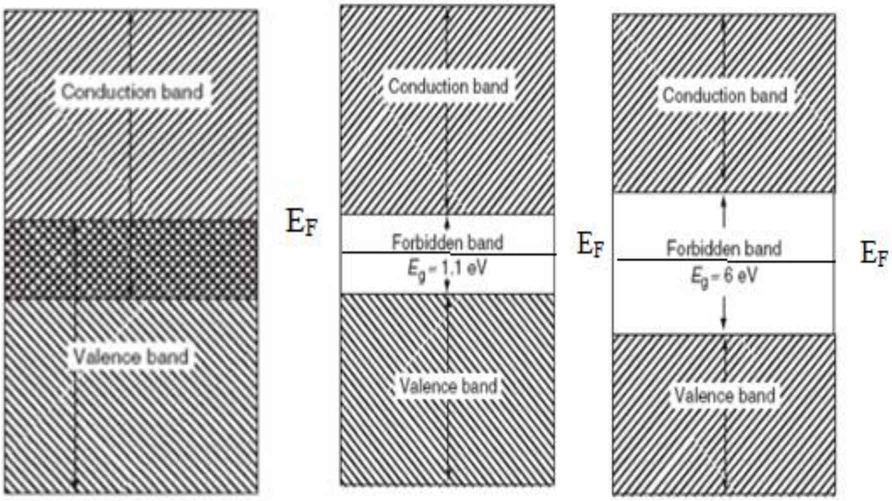
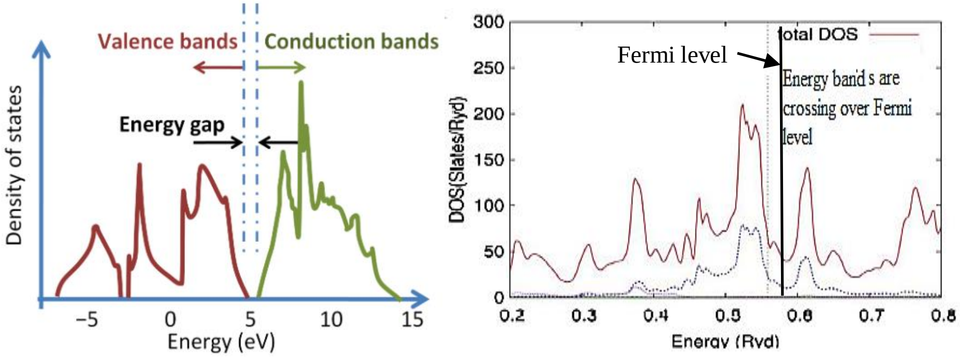


Fig. 2.5 (a) Conductor (b) Semiconductor (c) Insulator

Because in an n-type semiconductor will have a Fermi energy closes to the conduction band, whereas a p-type semiconductor will have Fermi energy close to the valence band. In metals this means that it gives us the velocity of the electrons during conduction. So during the conduction process, only electrons that have an energy that is close to that of the Fermi energy can be involved in the process.The concept of the Fermi energy is vital key to understanding electrical and thermal properties of solids. Another factor that the Fermi energy plays is in the role of understanding specific heat of solids at room temperatures.

The Fermi level describes the probability of electrons occupying a certain energy state, but in order to correctly associate the energy level the number of available energy states need to be determined is known as **Density of States (N (E_F)).** It will be useful to study about the conducting and physical properties of the materials. It gives the number of allowed electron (or) hole states per volume at a given energy. In the energy vs Density of States profile if the electron states are crosses the Fermi level are considered as metals or else it is considered as insulators [refer section 2.4.4 and 2.4.5].

Example:



Density of states (DoS) histogram for insulator (Left) Conductor (Right)

Fermi Surface

At absolute zero, the Fermi surface (constant energy surface in k-space) is the highest filled energy surface in k or wave vector space (The k-space is the relation of particle energy (E) versus the momentum as a wave number (k)). Alternatively it can be defined as the boundary between the filled and empty states in ground state of the crystal. The Fermi surface is a useful tool to predict thermal, electrical, magnetic, and optical properties of materials.

Characteristics:

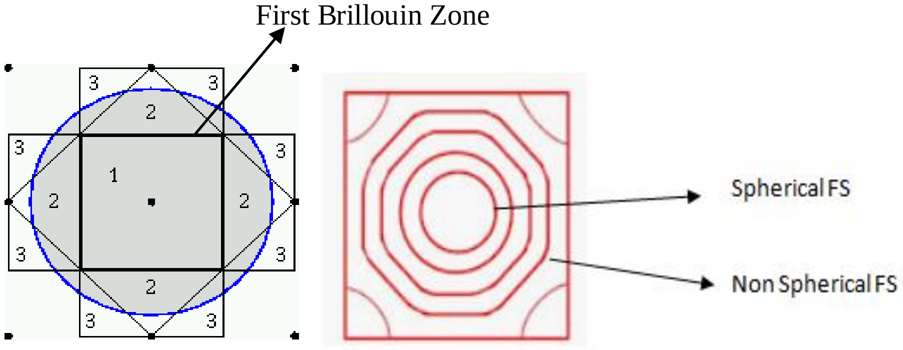


Fig.2.6 (a) Fermi surface (FS) in BrillouinZone (b) Shapes of FS

- The volume of the Fermi Surface represents the number of conduction electrons.

- The shape of the Fermi surface (Fig. 2.6) is derived from the periodicity and symmetry of the crystalline lattice and energy bands.

The Fermi surface is spherical within the first Brillouin zone, when the velocity (v) in k-space is proportional and parallel to the wave vector K.

The Fermi surface is non-spherical in higher zone, when the velocity is proportional to the gradient of the energy in k-space.

- The degree of distortion depends (a) how near the FS to the zone boundaries (b) The magnitude of the effective potential
- When an electrical field is applied, the Fermi surface shifts in the opposite direction of the field direction. This is shown below in Fig.2.7.Even though the electrons are cancelled out due to opposite direction of motion, some electrons remain un compensated resulting in a net current.

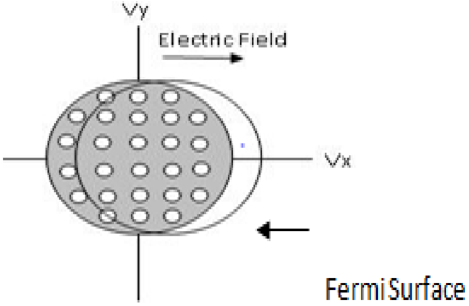


Fig.2.7 Due to electric field there is displacement in the Fermi surface

- It is used to know about the heat capacity, Pauli's paramgnetism and electrical conductivity etc.,

2.4.4 Semi metals

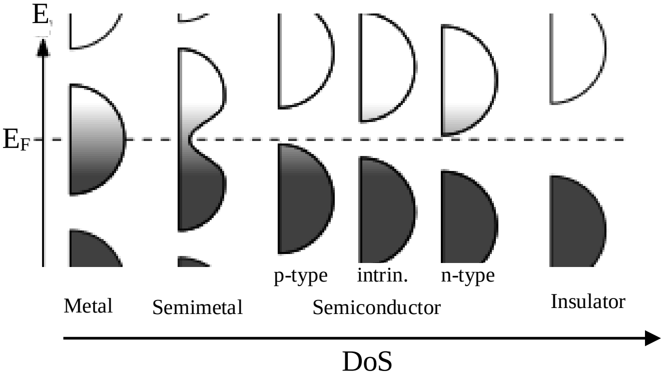


Fig. 2.8Comparison ofBand structure (Energy vsDoS)

One additional type of electron energy spectrum for solids is a "semimetal," two distinct energy bands overlap slightly at the Fermi level (shown in Fig.2.8). The semimetals are metals. Semimetals have properties intermediate between metals and semiconductors. Indeed they do allow conduction at absolute zero temperature. However, their further behavior is noticeably different from true metals because the overlap of the two bands is only small. One difference is that the electrical conduction of semimetals increases with temperature, unlike that of metals. Likely as semiconductors, semimetals at a higher temperature means that there are more electrons in the upper band and more holes in the lower band.

Examples: Arsenic, Antimony, Bismuth and a single sheet of carbon, called graphene, is right on the boundary between semimetal and semiconductor.

2.4.5 Half Metals

Half-metals were first described in 1983, as an explanation for the electrical properties of Mn-based Heusler alloys. Half-metals have attracted some interest for their potential use in spintronics.

A **half-metal** is any substance that acts as a conductor to electrons of one spin orientation, but as an insulator or semiconductor to those of the opposite orientation. Although all half-metals are ferromagnetic types but all ferromagnets are not half-metals. A ferromagnetic material with high spin polarization above room temperature is most important in the spintronics area.

In half-metals, the valence band for one spin orientation is partially filled while there is a gap in the density of states for the other spin orientation. This results in conducting behavior for only electrons in the first spin orientation. In some half-metals, the majority spin channel is the conducting while in others are in the minority channel.

The spin polarization (P) is defined as the ratio of the density of states (D) of up-spin and down-spin electrons at a Fermi level, $P=(D_{up}-D_{down})/(D_{up}+D_{down})$, as shown in Fig. 2.9. Since the density of states of up-spin and down-spin electrons are equal in paramagnetic materials, P=0 for paramagnetic materials. On the other hand, since the density of state of up-spin and down-spins are different in ferromagnetic materials, P is larger than 0, but smaller than 1 for ferromagnetic materials.

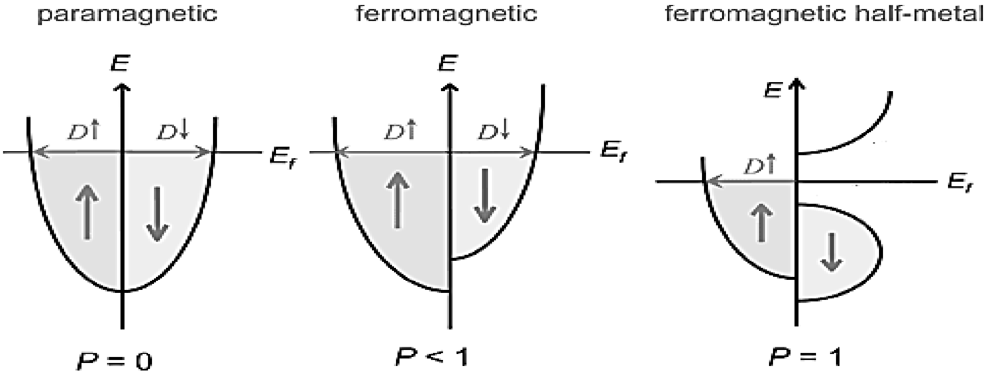


Fig. 2.9Density of states of paramagnetic, ferromagnetic and ferromagnetic half-metal materials

The P values of Fe and Co are known to be around 0.5. If a material has a band gap in the minority band (semiconducting) at a Fermi level and exhibits metallic behavior in the majority band, the density of state of the minority band is zero at the Fermi level. In this case, since only up-spin electrons are present at the Fermi level, P=1. This type of material is called "half-metal" because it exhibits both metallic and semiconducting behaviors.

Examples: chromium (IV) oxide, magnetite, lanthanum strontium manganite (LSMO), chromium arsenide. Many of the known examples of half-metals are oxides, sulfides, or Heusler alloys.

2.5 BLOCH THEOREM

Based on band theory, the solids are again classified into conductor, semiconductor and insulator. The free electron theory explained about the electrical conductivity and thermal conductivity of solids but it fails to clearly explain about the difference in between the conductor, insulator and semiconductor. Hence the Bloch theorem is used to explain as follows:

In order to understand the difference between the conductors and insulators, it is necessary to incorporate the variation of potential inside the crystal due to the presence of positive ion cores in the free electron model. It gives the realistic idea about to assume the potential inside the metallic crystal to be periodic with the periodicity of the lattice shown in Fig. 2.10.

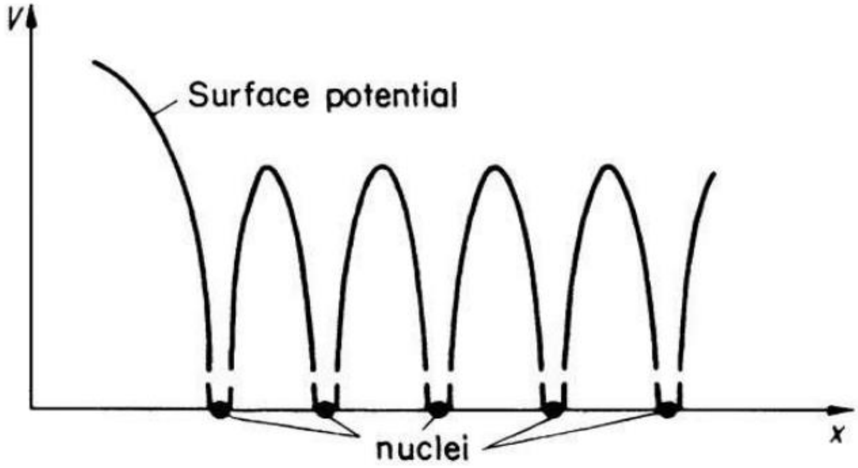


Fig. 2.10 One dimensional periodic potential in a crystal

Suppose an electron passes along X-direction in a one-dimensional crystal having periodic potentials:

$$V(x) = V(x + a)$$

where ‘a’ is the periodicity of the potential.

The Schrödinger wave equation for the moving electron in a constant potential (V) is:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}[E - V(x)]\psi = 0 \qquad (1)$$

The solution for Equation (1) is of the form:

$$\psi(x) = e^{iKx}\, u_k(x) \qquad (2)$$

where $$u_k(x) = u_k(x + a) \qquad (3)$$

Eqn. (3) represents periodic function and e^{ikx} represents plane wave. The above statement is known as Bloch theorem and Eqn. (3) is called **Block function.**

The Bloch function has the property:

$$\psi(x + a) = \exp\,[ik\,(x + a)]\; u_k(x + a) = \psi(x) \exp ika \qquad (4)$$

or $\psi(x + a) = Q\psi\,(x)$ where $Q = \exp ika$ (5)

Proof: Suppose g(x) and f(x) are two real independent solutions for the Schrödinger's equation (1), then f($x + a$) and g($x + a$) are also solutions of the above equation so that:

$$f(x + a) = \alpha_1 f(x) + \alpha_2 g(x)$$

$$g(x + a) = \beta_1 f(x) + \beta_2 g(x) \qquad (6)$$

where α_1, α_2, β_1 and β_2 are the real functions of energy (E).

The solution for Schrödinger wave equation is of the form:

$$\psi(x) = \mathrm{A}\, f(x) + \mathrm{B}\, g(x) \qquad (7)$$

where A and B are constants and

$$\psi(x + a) = \mathrm{A}\, f(x + a) + \mathrm{B}\, g(x + a) \qquad (8)$$

Substituting Eqn (6) in Eqn (8), we have:

$$\psi(x + a) = \mathrm{A}[\alpha_1 f(x) + \alpha_2 g(x)] + \mathrm{B}[\beta_1 f(x) + \beta_2 g(x)]$$

$$= [\mathrm{A}\alpha_1 + \mathrm{B}\beta_1]\, f(x) + [\mathrm{A}\alpha_2 + \mathrm{B}\beta_2]\, g(x) \qquad (9)$$

From the property of Bloch function, Eqn. (5) can be written as using Eqn. (7), we have:

$$\psi(x + a) = Q\psi(x) = Q\mathrm{A}f(x) + Q\mathrm{B}g(x) \tag{10}$$

Comparing Equations (9) and (10), we have:

$\mathrm{A}\alpha_1 + \mathrm{B}\beta_1 = Q\mathrm{A}$

and $\mathrm{A}\alpha_2 + \mathrm{B}\beta_2 = Q\mathrm{B}$ (11)

In Eqn. (11), A and B have non-vanishing values only if the determinant of the coefficients of A and B is equal to zero.

i.e. $$\begin{vmatrix} \alpha_1 - Q & \beta_1 \\ \alpha_2 & \beta_2 - Q \end{vmatrix} = 0$$

or $$Q^2 - (\alpha_1 + \beta_2)\, Q + \alpha_1\beta_2 - \alpha_2\beta_1 = 0 \tag{12}$$

From the above equation, we can show $\alpha_1\beta_2 - \alpha_2\beta_1 = 1$ Then the Eqn.(12) becomes:

$$Q^2 - (\alpha_1 + \beta_2)Q + 1 = 0 \tag{13}$$

The above quadratic equation has two roots say Q_1 and Q_2. So, we have two values for $\psi(x)$ i.e., $\psi_1(x)$ and $\psi_2(x)$. Also note that $Q_1 Q_2 = 1$. For certain values of energy corresponding to $(\alpha_1 + \beta_2)^2 < 4$, the two roots are complex and can be written as:

$$Q_1 = e^{iKa} \quad \text{and} \quad Q_2 = e^{-iKa} \tag{14}$$

The wave functions $\psi_1(x)$ and $\psi_2(x)$ can be represented as $\psi_1(x + a) = e^{ika}\, \psi_1(x)$ and

$$\psi_2(x + a) = e^{-iKa}\, \psi_2(x) \tag{15}$$

For other regions of energy corresponding to $(\alpha_1 + \beta_2)^2 > 4$, the roots Q_1 and Q_2 are real and reciprocal to each other. These two roots corresponding to Schrödinger's equation of the type:

$$\psi_1(x) = e^{\mu x}u(x) \quad \text{and} \quad \psi_2(x) = e^{-\mu x}u(x) \tag{16}$$

(μ is a real quantity)

Mathematically, the above solutions are sound but not accepted as wave functions describing elections. This leads to the energy spectrum of an electron in a periodic potential consists of allowed and forbidden energy regions or bands.

2.6 SEMICONDUCTING MATERIALS

2.6.1Classification of Semiconductors based on Energy band theory

Semiconductors are classified into two types. Intrinsic semiconductors and Extrinsic semiconductors.

a) Intrinsic Semiconductors: A semiconductor in an extremely pure form is known as Intrinsic semiconductor. The vacant conduction band is separated by a forbidden energy gap E_g from the filled valence band. So at absolute zero (0^0k), there is no electric conduction. Fermi level is at the middle of valance band and conduction band.

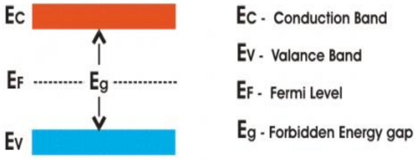


Fig. 2.11. Intrinsic semiconductor

b) Extrinsic Semiconductors: When an impurity is added to an intrinsic semiconductor its conductivity changes. This process of adding impurity to a semiconductor is called Doping and the impure semiconductor is called extrinsic semiconductor. Depending on the type of impurity added, extrinsic semiconductors are further classified as n-type and p-type semiconductor.

n-type semiconductor

When an impurity (electron) is added to an intrinsic semiconductor its conductivity changes and is called n-type semiconductor. This extra electron is forming a new energy level called **donor energy level** and is situated just below the conduction band. Fermi level is at the middle of donor energy level and conduction band.

p-type semiconductor

When an impurity (hole) is added to an intrinsic semiconductor its conductivity changes and is called p-type semiconductor. This extra hole is forming a new energy level called **acceptor energy level** and is situated just above the valance band. Fermi level is at the middle of acceptor energy level and conduction band.

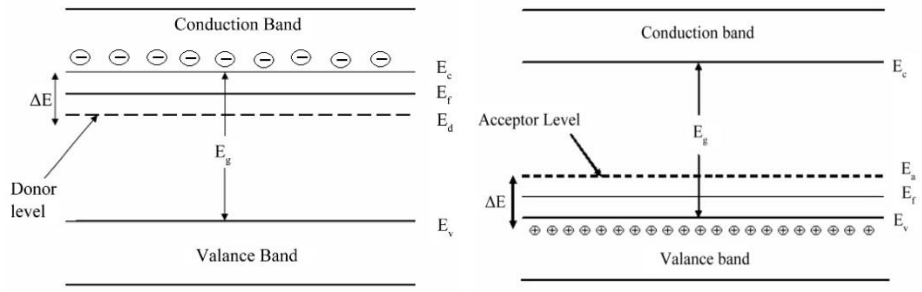


Fig. 2.12 (a) n-type semiconductor (b) p-type semiconductor

The detailed description about intrinsic and extrinsic semiconductor as follows

(a)Intrinsic or Pure Semiconductor:

A semiconductor in an extremely pure form is known as Intrinsic semiconductor. Example: Silicon, germanium. Both silicon and Germanium are tetravalent (having 4 valence electrons). Each atom forms a covalent bond or electron pair bond with the electrons of neighbouring atom. The structure is shown below.

At absolute zero temperature 0^0 k, all the valence electrons are tightly bounded the nucleus hence no free electrons are available for conduction. The semiconductor therefore behaves as an Insulator at absolute zero temperature. The vacant conduction band is separated by a forbidden energy gap E_g from the filled valence band.

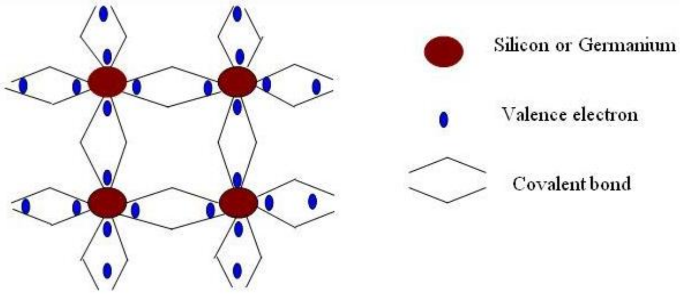


Fig. 2.13Intrinsic semiconductor

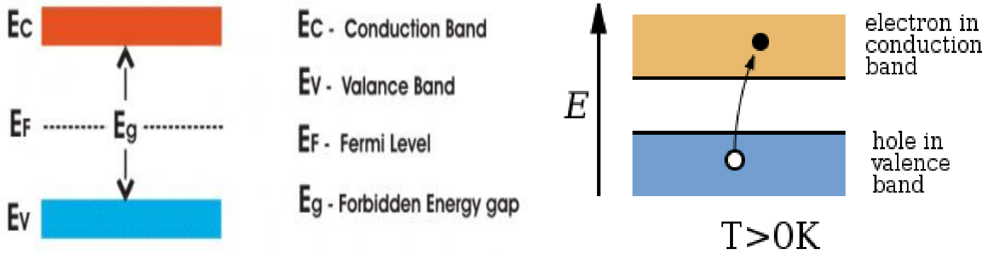


Fig. 2.14(a) Band Structure of an Intrinsic Semiconductor at 0^{o}K (b) Intrinsic Semiconductor – Holes and Electrons

When the temperature is increased, the electrons are thermally excited from the valence band to the conduction band. So they (electrons) become free in the conduction band. Therefore, electric conduction is possible in the conduction band. Due to this transition, the number of electrons in the conduction band will be more, but the same numbers of electrons are missing in the valence band. This gives the same number of 'holes' in the valence band as shown in figure. These holes act as a positive charge carrier. The electrons in the conduction band and holes in the valence band are responsible for electric conduction.

When a hole is created in the valence band, the electron from the nearby covalent bond (in the valence band) breaks its own bond and moves forward to fill the previous hole. Therefore, a new covalent bond is again formed in the valance band and this createsanother hole in the valence band. This hole is considered as particle similar to the electron and it have positive charge with definite mass. In the conduction band conduction is possible by electrons, since the electrons in the conduction band move freely inside the crystal. In the valence band, conduction is possible by holes since the movement of electrons in the valence band is restricted.

(b) Extrinsic semiconductors: When an impurity is added to an intrinsic semiconductor its conductivity changes. This process of adding impurity to a semiconductor is called Doping and the impure semiconductor is called extrinsic semiconductor. Depending on the type of impurity added, extrinsic semiconductors are further classified as n-type and p-type semiconductor.

n-type semiconductor: When a small amount of pentavalent impurity is added to a pure semiconductor (extrinsic semiconductor) is called as **n-type semiconductor.** Addition of Pentavalent impurity provides a large number of free electrons in a semiconductor crystal. Typical example for pentavalent impurities is Arsenic, Antimony and Phosphorus etc. Such impurities which produce n-type semiconductors are known as Donor impurities because they donate or provide free electrons to the semiconductor crystal.

To understand the formation of n-type semiconductor, consider a pure silicon crystal with impurity say arsenic added to it. We know that a silicon atom has 4 valence electrons and Arsenic has 5 valence electrons. When Arsenic is added as impurity to silicon, the 4 valence electrons of silicon make co-valent bond with covalent bond 4 valence electrons of Arsenic. The 5th Valence electrons finds no place in the thus, it becomes free and travels to the conduction band as shown in figure. Therefore, for each arsenic atom added, one free electron will be available in the silicon crystal. Though each arsenic atom provides one free electrons yet an extremely small amount of arsenic impurity provides enough atoms to supply millions of free electrons.

Due to thermal energy, still hole election pairs are generated but the number of free electrons are very large in number when compared to holes. So in an n-type semiconductor electrons are majority charge carriers and holes are minority charge carriers. Since the current conduction is pre-dominantly by free electrons (-vely charges) it is called as n-type semiconductor (n- means –ve).

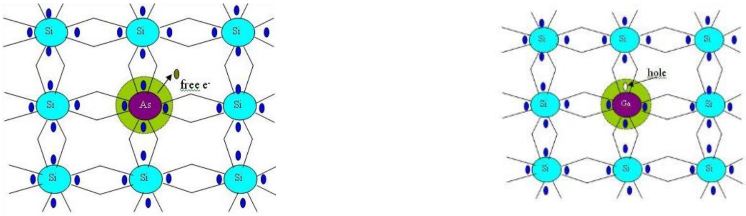


Fig. 2.15(a) n-type semiconductor (b) p-type semiconductor

p-type semiconductor: When a small amount of trivalent impurity is added to a pure semiconductor it is called p-type semiconductor**.** The addition of trivalent impurity provides large number of holes in the valance band in semiconductor crystals. Example: Gallium, Indium or Boron etc. Such impurities which produce p-type semiconductors are known as acceptor impurities because the holes created can accept the electrons in the semiconductor crystal.

To understand the formation of p-type semiconductor, consider a pure silicon crystal with an impurity say gallium added to it. We know that silicon atom has 4 valence electrons and Gallium has 3 electrons. When Gallium is added as impurity to silicon, the 3 valence electrons of gallium make 3 covalent bonds with 3 valence electrons of silicon. The 4th valence electrons of silicon cannot make a covalent bond with that of Gallium because of short of one electron as shown above. This absence of electron is called a hole. Therefore for each gallium atom added one hole is created, a small amount of Gallium provides millions of holes.Due to thermal energy, still hole-electron pairs are generated but the number of holes is very large compared to the number of electrons. Therefore, in a p-type semiconductor holes are majority carriers and electrons are minority carriers. Since the current conduction is predominantly by hole (+ charges) it p-type semiconductor (p means +ve).

2.7. CARRIER (ELECTRON AND HOLE) CONCENTRATION OF INTRINSIC SEMICONDUCTOR

At absolute zero temperature, the valence band is completely occupies by electrons but the conduction band is completely empty. The completely filled band does not conduct current. Thus, at low temperatures, the pure (intrinsic) semiconductor gives an extremely high resistance to electric conduction.

At finite temperature, the electrons in the valence band get excited and they will move to the conduction band. So holes are produced in the valence band. The probability of occupation of electrons in the conduction band and holes in the valence band are described by the Fermi distribution function .i.e..,

$$\text{f(E)} = \frac{1}{1+exp\left(\frac{E-E_F}{K_BT}\right)} \qquad (1)$$

where, E_F is the Fermi Energy

K_B is the Boltzman's constant

T is the temperature of the intrinsic semiconductor

The density of states of conduction and valence bands.

Electron Concentration of Intrinsic Semiconductor

Generally, the band gap ($E-E_F$) is of the order of 1 eV. For electrons at the bottom of the conduction band, $exp\frac{(E-E_F)}{K_{B-T}}=e^{20} \gg 1$. For other electrons in the conductions band, this factor will be still higher. Thus, the value +1 in the denominator is negligible. This approximation is called **Boltzman approximation.**

$$f(E) \sim \frac{1}{exp\left(\frac{E-E_F}{K_BT}\right)} = \exp\left(\frac{E_F-E}{K_BT}\right) \qquad (2)$$

The concentration of electron in the conduction band is

$$n_e = n = \int_{E_c}^{\infty} f(E)N_e(E)dE \qquad (3)$$

The concentration of holes in the valence band is,

$$n_n = p = \int_{-\infty}^{E_v}\left(1-f(E)\right) N_h(E)dE \qquad (4)$$

Where, $N_e(E)$ is the electron density (number of electrons per unit volume) of states near the conduction band and

$N_h(E)$is the hole density of states near the valence band.

For an intrinsic semiconductor, n = p.

The electron density of states in the conduction band is

$$N_e(E) = \frac{4\pi}{h^3}(2m_e^*)^{3/2}(E-E_c)^{1/2} \qquad (5)$$

Where $m_e^*$is the effective mass of electron in conduction band.

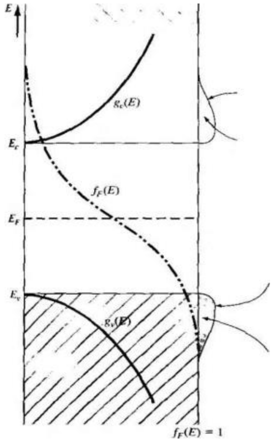


Fig. 2.16Density of states of conduction and valence bands of Intrinsic Semiconductor

Substituting equations (2) and (5) in equation (3), we get,

$$n_e = \frac{4\pi}{h^3}(2m_e^*)^{\frac{3}{2}}\int_{E_c}^{\infty}(E-E_c)^{\frac{1}{2}} \;.\; \exp\left(\frac{E_C-E}{K_BT}\right) . dE \qquad (6)$$

Let $\exp\left(\frac{E_F-E}{K_BT}\right) = exp\left(\frac{E_F-E_C+E_C-E}{K_BT}\right)$

$$= exp\left(\frac{E_F-E_C}{K_BT}\right) . exp\left(\frac{E_C-E}{K_BT}\right)$$

Substituting this in equation (6), we get,

$$n_e = \frac{4\pi}{h^3}(2m_e^*)^{\frac{3}{2}}\int_{E_c}^{\infty}(E-E_c)^{\frac{1}{2}} . exp\left(\frac{E_{F-E_C}}{K_BT}\right) . \exp\left(\frac{E_C-E}{K_BT}\right) . dE \qquad (7)$$

Let$\left(\frac{E-E_c}{K_BT}\right)$ =x   Therefore $E-E_c$=x.K_BT

Differentiating the above equation, we get,   dE =K_BT.dx

Also$(E-E_c)^{\frac{1}{2}}$=$x^{1/2}.K_B{}^{1/2}T^{1/2}$

When   E =E_c,x = 0   and when   E =∞, $x=\infty$

Substituting the above values in equation (7), we get,

$$n_e = \frac{4\pi}{h^3}(2m_e{}^*)^{3/2}\int_0^\infty x^{\frac{1}{2}}K_B{}^{1/2}T^{1/2}.\exp\left(\frac{E_F-E_c}{K_BT}\right)e^{-x}.K_BT.dx$$

$$=\frac{4\pi}{h^3}(2m_h{}^*K_BT)^{3/2}\,.exp\left(\frac{E_F-E_c}{K_BT}\right)\int_0^\infty x^{1/2}\ e^{-x}.dx$$

Now$\int_0^\infty x^{1/2}\ e^{-x}dx = \frac{\pi^{1/2}}{2}$

$$n_e = \frac{4\pi}{h^3}(2m_e{}^*K_BT)^{3/2}.exp\left(\frac{E_F-E_C}{K_BT}\right).\left(\frac{\pi^{1/2}}{2}\right)$$

$$n_e = 2\left(\frac{2\pi m_e{}^*K_BT}{h^2}\right)\exp\left(\frac{E_F-E_C}{K_BT}\right) \qquad (8)$$

This is the concentration of electron in the conduction band of intrinsic semiconductor.

2.7.1Hole Concentration of Intrinsic Semiconductor

For holes at the top of the valence band, exp$\left(\frac{E-E_F}{K_BT}\right) \ll 1$. For other holes in the valence band this factor will be still lower. Now,

$$[1-f(E)] = 1-\frac{1}{1+exp\left(\frac{E-E_F}{K_BT}\right)}$$

$$= \frac{1+exp\left(\frac{E-E_F}{K_BT}\right)-1}{1+exp\left(\frac{E-E_F}{K_BT}\right)} \qquad = \frac{exp\left(\frac{E-E_F}{K_BT}\right)}{1+exp\left(\frac{E-E_F}{K_BT}\right)}$$

The exponential factor in the denominator is negligible, since$exp\left(\frac{E-E_F}{K_BT}\right) \ll 1$.

$$[1-f(E)] \cong exp\left(\frac{E-E_F}{K_BT}\right) \qquad (9)$$

The hole density of states in the valence band is,

$$N_h(E) = \frac{4\pi}{h^3}(2m_h{}^*)^{3/2}\ (E_v-E)^{1/2} \qquad (10)$$

Where $m_h{}^*$ is the effective mass of hole in the valence band.

Substituting equation (9)and(10)in equation (4), we get,

$$n_h = \frac{4\pi}{h^3}(2m_h{}^*)^{\frac{3}{2}}\int_{-\infty}^{E_v} (E_v-E)^{\frac{1}{2}} exp\left(\frac{E-E_F}{K_BT}\right).dE \qquad (11)$$

Let $exp\left(\frac{E-E_F}{K_BT}\right) = exp\left(\frac{E-E_v+E_v-E_F}{K_BT}\right)$

$$= exp\left(\frac{E-E_V}{K_BT}\right).\exp\left(\frac{E_v-E_F}{K_BT}\right)$$

Substituting this in equation (11), we get,

$$n_{h} = P = \frac{4\pi}{h^3}(2m_h{}^*)^{\frac{3}{2}}\int_{-\infty}^{E_v} (E_v-E)^{\frac{1}{2}} exp\left(\frac{E-E_v}{K_BT}\right).\exp\left(\frac{E_v-E_F}{K_BT}\right).dE \quad (12)$$

Let $\frac{E_v-E}{K_BT}$=x $\qquad E_v - E = x.K_BT$

Differentiating the above equation, we get, -dE $=K_BT.dx$

dE $=-K_BT.dx$

$$(E_v-E)^{1/2} = x^{1/2}K_B{}^{1/2}T^{1/2}$$

When E=$-\infty$, $x = \infty$ and when E=E_v, x=0

Substituting the above values in equation (12), we get

$$n_{h\,=\,P\,=}\frac{4\pi}{h^3}(2m_h{}^*)^{\frac{3}{2}}\int_{\infty}^{0} x^{\frac{1}{2}}K_B{}^{\frac{1}{2}}T^{\frac{1}{2}}e^{-x}exp\left(\frac{E_{v-E_F}}{K_BT}\right).(-K_BT)dx$$

$$=\frac{-4\pi}{h^3}(2m_e{}^*K_BT)^{\frac{3}{2}}exp\left(\frac{E_{v-E_F}}{K_BT}\right)\int_{\infty}^{0} x^{\frac{1}{2}}e^{-x}.dx$$

Now $\int_{\infty}^{0} x^{\frac{1}{2}}e^{-x}.dx = -\int_{0}^{\infty} x^{\frac{1}{2}}e^{-x}.dx = -\frac{\pi^{\frac{1}{2}}}{2}$

$$n_h = \frac{-4\pi}{h^3}(2m_h{}^*K_BT)^{3/2}.exp\left(\frac{E_{v-E_F}}{K_BT}\right).(-\frac{\pi}{2}^{1/2})$$

$$n_h = 2\left(\frac{2\pi m_h{}^*K_BT}{h^2}\right)^{3/2}.\exp\left(\frac{E_{v-E_F}}{K_BT}\right) \qquad (13)$$

This is the concentration of hole in the valence bandof intrinsic semiconductor.

2.7.2Fermi Energy of Intrinsic Semiconductor

For an intrinsic semiconductor, $n_e = n_h(i.e, \quad n = p)$

$$2\left(\frac{2\pi m_e{}^*K_BT}{h^2}\right)^{3/2}.\exp\left(\frac{E_{F-E_C}}{K_BT}\right) = 2\left(\frac{2\pi m_h{}^*K_BT}{h^2}\right)^{3/2}.\exp\left(\frac{E_{v-E_F}}{K_BT}\right)$$

$$(m_e{}^*)^{3/2}\exp\left(\frac{E_{F-E_C}}{K_BT}\right)=(m_h{}^*)^{3/2}\exp\left(\frac{E_{v-E_F}}{K_BT}\right)$$

Taking logarithm on both sides, we get,

$$\log(m_e{}^*)^{3/2}+\left(\frac{E_{F-E_C}}{K_BT}\right) = \log(m_h{}^*)^{\frac{3}{2}}+\left(\frac{E_{v-E_F}}{K_BT}\right)$$

$$\left(\frac{E_{F-E_C}}{K_BT}\right)-\left(\frac{E_{v-E_F}}{K_BT}\right) = \frac{3}{2}log(m_h{}^*) - \frac{3}{2}log(m_e{}^*)$$

$$\frac{2E_F-(E_c+E_v)}{K_BT}=\frac{3}{2}log\left(\frac{m_h{}^*}{m_e{}^*}\right)$$

$$2E_F-(E_c+E_v)=\frac{3}{2}K_BTlog\left(\frac{m_h{}^*}{m_e{}^*}\right)$$

$$2E_F=(E_c+E_v)+\frac{3}{2}K_BTlog\left(\frac{m_h{}^*}{m_e{}^*}\right)$$

$$E_F=\left(\frac{E_c+E_v}{2}\right)+\frac{3}{4}K_BTlog\left(\frac{m_h{}^*}{m_e{}^*}\right) \qquad (14)$$

When $m_e{}^*=m_h{}^*$,

Then $\log\left(\frac{m_h{}^*}{m_e{}^*}\right)=\log 1=0$,

$$\therefore E_F=\frac{E_c+E_v}{2}$$

That is, Fermi level lies exactly half the distance between the top of the valence band and bottom of the conduction band. When the temperature of an intrinsic semiconductor increase, $m_h{}^*>m_e{}^*$and the Fermi level is raised.

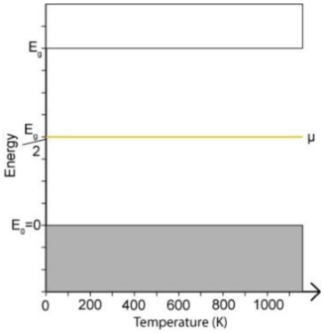


Fig. 2.17. Fermi Energy of Intrinsic Semiconductor

2.8CARRIER (ELECTRON AND HOLE) CONCENTRATION OF INTRINSIC SEMICONDUCTOR IN TERMS OF BAND GAP ENERGY

Then from equation (14),

$$E_F-E_c=\left(\frac{E_c+E_v}{2}\right)+\frac{3}{4}K_BTlog\left(\frac{m_h{}^*}{m_e{}^*}\right)-E_c=\left(\frac{E_v-E_c}{2}\right)+\frac{3}{4}K_BTlog\left(\frac{m_h{}^*}{m_e{}^*}\right)$$

The above equation is divided by $K_BT.we$ get,

$$\left(\frac{E_F-E_c}{K_BT}\right)=\frac{-E_g}{2K_BT}+\frac{3}{4}log\left(\frac{m_h{}^*}{m_e{}^*}\right) \qquad (15)$$

Where $E_g = E_c - E_v$ is the energy of the forbidden gap.

Substituting equation (15) in equation (8), we get,

$$n_e = n = 2\left(\frac{2\pi m_e{}^* K_B T}{h^2}\right)^{3/2} \exp\left[\frac{-E_g}{2K_B T} + \frac{3}{4} log\left(\frac{m_h{}^*}{m_e{}^*}\right)\right]$$

$$n_e = n = 2\left(\frac{2\pi m_e{}^* K_B T}{h^2}\right)^{3/2} \exp\left[\frac{-E_g}{2K_B T}\right]\exp\left[log\left(\frac{m_h{}^*}{m_e{}^*}\right)^{3/4}\right]$$

$$n_e = n = 2\ \left(\frac{2\pi m_e{}^* K_B T}{h^2}\right)^{3/2}\left(\frac{m_h{}^*}{m_e{}^*}\right)^{3/4} .exp\left(\frac{-E_g}{2K_B T}\right)$$

$$n_e = n = 2\ \left(\frac{2\pi K_B T}{h^2}\right)^{3/2} (m_e{}^* m_h{}^*)^{3/4}\ exp\left(\frac{-E_g}{2K_B T}\right) \qquad (16)$$

Also from equation(14),

$$E_v - E_F = E_v - \left[\frac{E_c + E_v}{2} + \frac{3}{4} K_B T log\left(\frac{m_h{}^*}{m_e{}^*}\right)\right]$$

$$= \left(\frac{E_v - E_c}{2}\right) - \frac{3}{4} K_B T log\left(\frac{m_h{}^*}{m_e{}^*}\right)$$

The above equation is divided by $K_B T$, weget,

$$\frac{E_{v-}E_F}{K_B T} = \frac{-E_g}{2K_B T} - \frac{3}{4} log\left(\frac{m_h{}^*}{m_e{}^*}\right) \qquad (17)$$

Substituting equation (17) in equation (13), we get,

$$n_h = p = 2\ \left(\frac{2\pi m_h{}^* K_B T}{h^2}\right)^{3/2} .exp\left[\frac{-E_g}{2K_B T} - \frac{3}{4} log\left(\frac{m_h{}^*}{m_e{}^*}\right)\right]$$

$$n_h = p = 2\ \left(\frac{2\pi m_h{}^* K_B T}{h^2}\right)^{3/2} .exp\left[\frac{-E_g}{2K_B T}\right]\exp\left[log\left(\frac{m_h{}^*}{m_e{}^*}\right)^{3/4}\right]$$

$$n_h = p = 2\left(\frac{2\pi m_h{}^* K_B T}{h^2}\right)^{\frac{3}{2}}\left(\frac{m_h{}^*}{m_e{}^*}\right)^{\frac{3}{4}} \exp\left(\frac{-E_g}{2K_B T}\right)$$

$$n_h = p = 2\left(\frac{2\pi K_B T}{h^2}\right)^{3/2}(m_e{}^* m_h{}^*)^{3/4}\, exp\left(\frac{-E_g}{2K_B T}\right) \qquad (18)$$

From equation (16)and (18), we get,

$$n = p = 2\left(\frac{2\pi K_B T}{h^2}\right)^{3/2}(m_e{}^* m_h{}^*)^{3/4}\, exp\left(\frac{-E_g}{2K_B T}\right) \qquad (19)$$

Therefore, the concentration of electron in the conduction band and hole in the valence band depend upon the temperature of the intrinsic semiconductor.

2.9. EXPERIMENTAL DETERMINATION OF BAND GAP OF A SEMICONDUCTOR

The diode is immersed in an oil bath that in turn is kept in a heating mantle. A thermometer is also kept in the oil path such that its measure bulb is just at the height of diode. The diode is reversed biased with the help of DC voltage obtained from a DC power supply and the current that flows through the reverse diode is measured with a milliammeter. The Circuit diagram of this experiment is shown in the Fig.2.18(a).

The power supply is switched on and the voltage is adjusting say five volts. The current through the diode and the room temperature are noted. The power supply in switched off and the heating mantle is switched on. The oil bath is heated up to 65^0 C. The oil is stirred well. The temperature of oil bath stabilized say at 75^0 C.

The power supply is again switched on and the voltage is kept at 5V. The temperature (say 75^0C) and the corresponding current through the diode are noted.

Now the oil bath is allowed to cool slowly.As the temperature falls, the current through the diode decreases. The current falls for various temperatures (any convenient interval) are noted down.

The calculations are completed and a graph is plotted taking 1000/T on X-axis and log I_0 on Y-axis. A straight line is obtained as shown in the Fig.2.18 (b). The slow of the straight line is determined and using it the band gap is calculated using the following equation.

$$E_g = 3.973\times 10^{-4}\times \text{slopeeV}$$

Care should be taken while performing the experiment the diode and the thermometer are placed at the same level in the oil bath.The maximum temperature of the diode should not be beyond 80^0C. Reading of the current and temperature to be taken simultaneously.

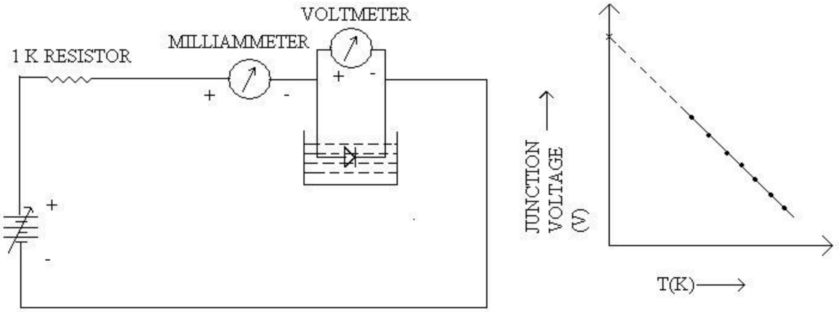


Fig. 2.18(a) Circuit diagram **(b) Expermental graph**

2.10 SUPERCONDUCTING MATERIALS

2.10.1 Introduction

The initial discovery of superconductive materials was made in the year 1911. Before 1986, the critical temperatures for all known superconductors did not exceed 23 Kelvin. Before the discovery and development of high temperature superconducting (HTSC) materials, the use of superconductivity had not been practical for widespread commercial applications, except for Magnetic Resonance Imaging (MRI) and Superconducting Magnetic Energy Storage (SMES) applications, principally because commercially available superconductors (i.e. low temperature superconducting (LTSC) materials) are made superconductive only when these materials are cooled to near 0K. Although it is technologically possible to cool LTSC materials to a temperature at which they become superconductive, broad commercialization of LTSC materials has been inhibited by the high cost associated with the cooling process. For example, liquid helium, which can be used to cool materials to about 4K, and which has been commonly used to cool LTSC materials, is expensive and relatively costly to maintain.

In 1986, a breakthrough in superconductivity occurred when two scientists, Dr. K. Alex Muller and Dr. J. Georg Bednorz, identified a ceramic oxide compound which was shown to be superconductive at 36K. This discovery earned them a Nobel Prize for Physics in 1987, which is one of four Nobel Prizes that have been awarded for work on superconductivity. A series of related ceramic oxide compounds which have higher critical temperatures were subsequently discovered.

2.11 SUPERCONDUCTING MATERIALS AND TRANSITION TEMPERATURE (T_c)

The electrical resistivity of pure metals decreases gradually with decrease in temperature and reaches a low but measurable resistivity values at temperatures nearing zero Kelvin. However there are few materials for which the electrical resistivity at a very low temperature abruptly plunges from a finite value to zero and remains there upon further cooling. This phenomenon is called Superconductivity. Materials that display this behaviour are called superconducting materials. The temperature at which there is abrupt drop in resistivity to zero value is called as superconducting transition temperature (T_c). At T_c, the materials get transformed from one conducting state to another. Thus zero resistivity ($\rho = 0$) or infinite conductivity is observed in a superconductor at all temperatures below a critical temperature ($\rho = 0$ for all $T < T_c$).

At Transition temperature T_c,

1. The electrical resistivity of the material abruptly drops to zero.
2. The magnetic flux lines are excluded from the material.
3. There is a discontinuous change in specific heat capacity.
4. There are small change in thermal conductivity and volume of the material.

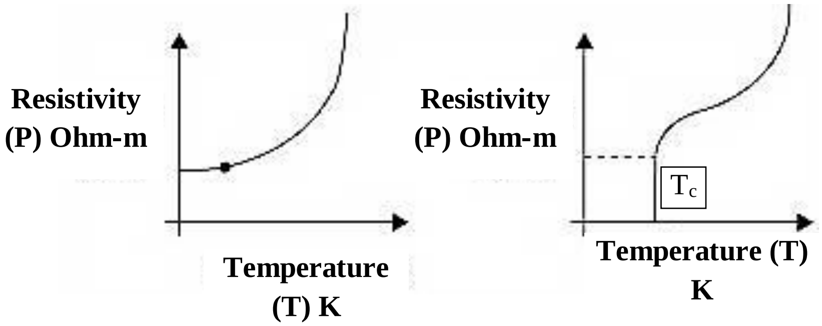


Fig. 2.19: Variation of resistivity with temperature (a) for normal conductors (b) for superconductors

2.12 OCCURRENCE OF SUPERCONDUCTIVITY

The occurrence of superconductivity was first discovered by Heike KamerlingOnnes in 1911. He liquefied Helium gas in 1908 whose boiling point is 4.2 K. He studied the properties of metals under low temperatures using liquid helium. In 1911, when he studied the electrical properties of solid mercury at very low temperatures, he found that the resistivity of mercury suddenly decreases 10^5 times around 4.2 K.

This drastic change in electrical resistivity of mercury indicates that mercury gets transformed from conducting state to superconducting state with the transition temperature of 4.2 K.

Since then, research is being aimed at getting the T_c nearer to room temperature. If the material is having low T_c (< 30 K) value, then the material is called low T_c superconductor. If the material is having high T_c (> 30 K) value, then the material is called high $T_c$superconductor.

The evolution of superconducting transition temperature (T_c) since its discovery.

Material	**Transition temperature (T_c) in K**	**Year**
Elements	**Type I Superconductors**	
Hg	4.2	1911
Pb	7.196	1913
Nb	9.46	1930
Compounds	**Type II Superconductors**	
NbN	16.0	1941
V_3Si	17.4	1950
Nb_3Sn	18.3	1956
Nb_3Ge	23.0	1935
Ceramics	**Type II Superconductors**	
La-Ba-CuOPerovskites	35.0	1986
Y-Ba-CuOPerovskites	92.0	1987
$Bi_2Sr_2Ca_2Cu_3O_{10}$	107	1988
$Bi_2Sr_2CaCu_2O_9$	110	1988
$Tl_2Ba_2Ca_2Cu_3O_{10}$	127	1988
$Hg_{12}Tl_3Ba_{30}Ca_{30}Cu_{45}O_{127}$	138	2006
$HgBa_2Ca_2Cu_2O_8$	153	2006
$(Tl_4Ba)Ba_4Ca_2Cu_{10}O_y$	240	2009

Note that the transition temperature of Y-Ba-CuOperovskites is 92K which is 15K above the boiling point of liquid nitrogen (77 K). The liquid nitrogen environment is far easier to handle and cheaper than liquid helium. Further it consumes about 25 times more energy to cool from 77 K to 4 K than from room temperature to 77 K.

2.13PROPERTIES OF SUPERCONDUCTING MATERIALS

2.13.1Zero Electrical Resistance

The DC electrical resistance of a superconductor at all temperatures below a critical temperature is practically zero. The transition from normal state to superconducting state occurs sharply in pure metals whereas it is not so in impure, HTSC oxides. In HTSC oxides, there exist several superconducting phases.

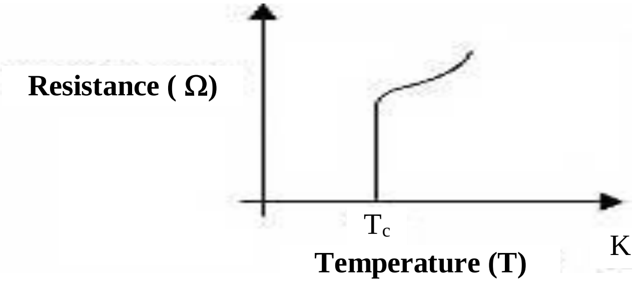


Fig. 2.20Variation of resistance with temperature

Consider a small amount of current applied to a superconducting ring. Since the ring is in superconducting state, it has zero resistance. Hence the superconducting current through the ring will keep on flowing without any appreciable change in its value. Such a current is said to be persistent current. e.g. Dr.Collins experimentally found that the resistance of a superconducting ring by passing standard current through it. $[(R_{scs} / R_{ns}) < 10^{-5}]$ where R_{scs} is the resistance of the material in superconducting state and R_{ns} is the resistance of the material in normal state. He observed that, the value of the current in the superconducting ring does not change for more than 2.5 years. This shows that, the superconductor has virtually zero electrical resistance and they can conduct electricity without resistance.

2.13.2Diamagnetic Property (Meissner Effect):

When a weak magnetic field is applied to a superconducting material at a temperature below transition temperature ($T < T_c$), the magnetic flux lines are excluded from the material (**B** = 0). So the superconductor acts as an ideal diamagnet. The ejection of magnetic lines of force, when the superconducting material is cooled below its T_c in a weak magnetic field is called Meissner effect.

Generally, magnetic flux density inside the material = **B** = μ_0 (H+I)

Here, magnetic flux density inside the superconductor = **B** = 0

$$\mu_0 (H+I) = 0$$

μ_0 is the permeability of free space = constant, So $\mu_0 \neq 0$

Then (H+I) = 0 and H = - I

Magnetic susceptibility of the superconductor = $\chi_m = (I/H) = I/(-I) = (-1)$

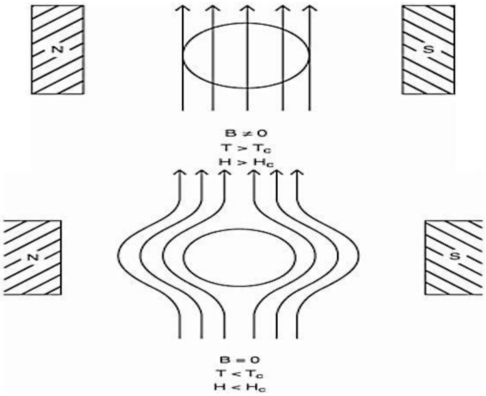


Fig. 2.21 (a) Magnetic field applied to a normal conductor (b) Magnetic field applied to a superconductor

For a diamagnetic material, χ_m is negative. So at T $<T_c$ and H $<H_c$, the superconducting material behaves like a **perfect diamagnetic material**.

2.13.3Effect of Heavy Magnetic Field:

Below T_c, superconductivity can be destroyed by the application of strong magnetic field. At any temperature, the minimum magnetic field required to destroy superconductivity is called critical field (H_c) of the material. Thus the value of $H_c$depends upon the temperature of the superconducting material and is given by $H_c = H_0 [1 - (T^2 / Tc^2)]$ where, H_0 is the critical field at 0 K, T is the temperature of the material and T_c is the transition temperature. From the graph, we know that, when the temperature of the material increases from 0 K, the value of Hc decreases and at T = T_c, H_c= 0. Also the value of H_c is different for different materials.

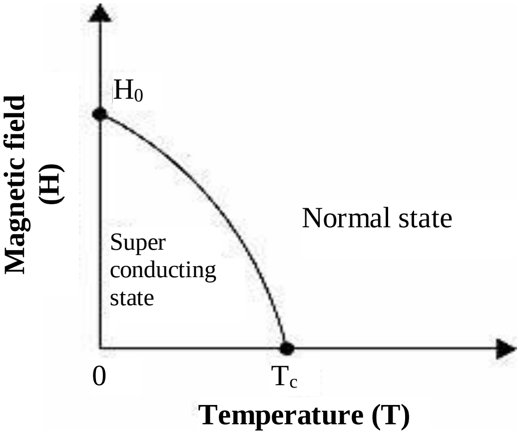


Fig. 2.22 Dependence of critical magnetic field with temperature

2.13.4Effect of Heavy Current (Silsbee's rule):

We know that a current carrying conductor has a magnetic field around it. When heavy current flows through a superconductor it will set up the magnetic field which destroys the superconducting property of the superconductor. According to

Silsbee’s rule, the critical current (I_c) flowing through a superconducting wire of radius r is $I_c = 2\pi r H_c$ where, H_c is the critical field at T K.

2.13.5Effect of Pressure:

Some materials which behave as normal conductors at normal pressure can undergo phase transition and behave as superconductors under the application of external pressure. *e.g.*

- Cesium becomes a superconductor at 110 kilo bar pressure and its $T_c = 1.5$ K
- Silicon becomes a superconductor at 165 kilo bar pressure and its $T_c = 8.3$ K

2.13.6Effect of Isotope:

The transition temperature of the given superconducting material varies with its mass number or isotopic mass and this effect is known as Isotope effect. Maxwell found that, T_c is inversely proportional to the square root of the atomic mass (M) of the isotope of the single superconductor. *i.e.*$T_c \propto [1 / M^{½}]$

For *e.g.* Mass number of mercury varies from 199.5 to 203.4 amu and the corresponding transition temperature $T_c$varies from 4.185 to 4.146 K respectively.

Obviously the existence of isotope effect indicates that although the superconductivity is an electronic phenomenon, it depends on the vibration of the crystal lattice (phonons) in which the electrons move.

2.13.7 Entropy:

Entropy is a measure of degree of disorderness of the given system. The entropy of a superconducting material is found to be lower than that of the normal conducting material which indicates that in the superconducting state the electrons are in the more ordered state than they are in the normal conducting state.

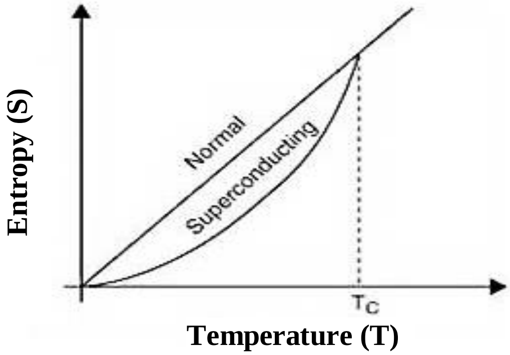


Fig. 2.23: Variation of entropy with temperature

2.13.8Specific heat capacity:

The specific heat capacity of a normal conductor (C_n) consists of two contributions, one from the conduction band electrons and the other from the lattice.

i.e. $C_n = C_n^{(electrons)} + C_n^{(lattice)}$; ‘n’ represents normal conductor.

i.e. $C_n = \sqrt{T} + \beta T^3$ where $\sqrt{T}$ arises from electron contribution and βT^3 arises from lattice contribution.

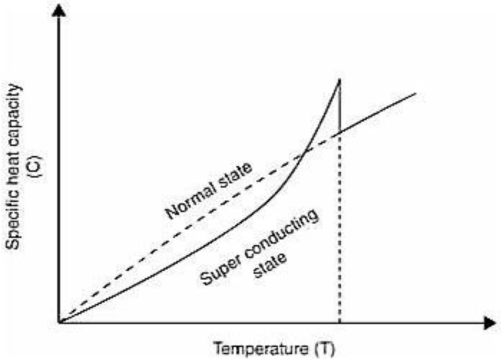
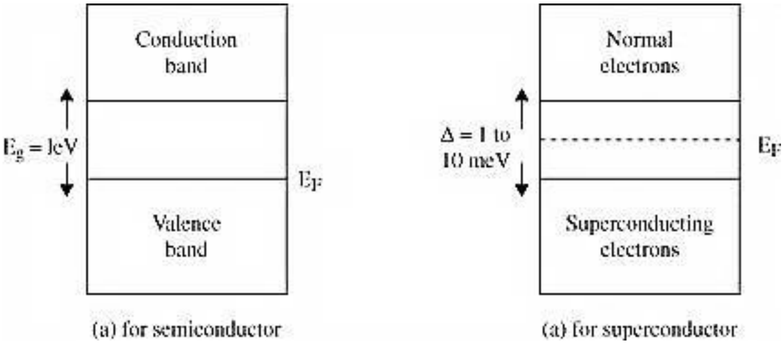


Fig. 2.24: Variation of Specific heat capacity with temperature

Consider the specific heat capacity of a superconductor (C_{es}) which varies in a characteristic way at T_c. It is given by $C_{es}(T < T_c) = A\ e^{-(\Delta / K_B T)}$ where K_B - Boltzmann constant, T- temperature of the specimen, Δ- a gap in the spectrum of allowed energy states separating the excited states from the ground states. From above equation, there appears a discontinuity in specific heat at T_c.

2.13.9 Energy gap:

Specific heat capacity measurements provided the first indication of an energy gap (Δ) in superconductors. The energy gap of a superconductor is of different nature than that of a semiconductor or a conductor. In case of semiconductors, the energy gap is tied to the lattice and it corresponds to the energy difference between the valence band and the conduction band (~1eV); whereas in superconductors, the energy gap (Δ) is tied to the Fermi energy and this gap in the spectrum of allowed energy states separates the excited states from the ground states. In conventional superconductors, Δ~1meV while in HTSC, Δ~1to 10 meV.



(a) for semiconductor (a) for superconductor

Fig. 2.25 Band Gap

2.13.10London Penetration Depth (λ)

When a superconducting material is placed in a magnetic field, a current is set up on the surface of the material and a magnetic field is produced everywhere within the material. This magnetic field opposes the applied external magnetic field.

So the applied magnetic field penetrates some distance into the superconductor. This distance is called London Penetration depth (λ). It is also defined as the effective depth to which a magnetic field penetrates a superconductor.

$H_{(x)} = H_o\, e^{-(x/\lambda)}$ where $H_{(x)}$ and H_o are the magnetic fields at a distance 'x' and on the surface. $\lambda = 10^3$ to 10^4 Å; λ is independent of the frequency of the magnetic field but dependent on temperature.

$(\lambda_T / \lambda_o)^2 = 1 / [1 - (T/T_c)^4]$ where λ_T and λ_o are the penetration depth at T K and 0 K.

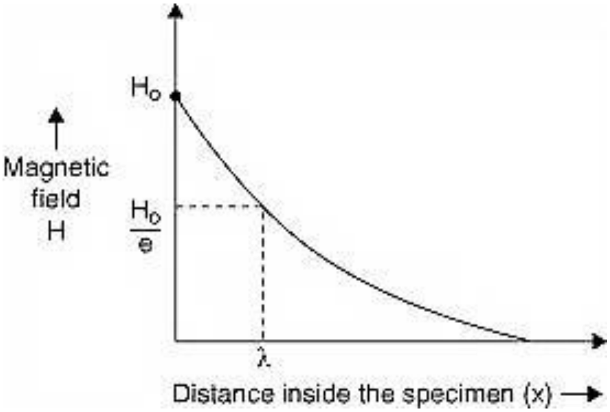


Fig. 2.26Penetration depth

2.13.11Coherence Length (ξ):

The paired electrons (Cooper pairs) are not scattered because of their peculiar properties of smoothly riding imperfection. Also there is no exchange of energy between them. Mean time, they can maintain their coupled motion up to certain distance (about 10^{-6} m). This distance is called Coherence Length. It is also refereed as the distance between two electrons of the Cooper pair within the highly coherent superconducting state and is given by $\xi = (\hbar V_F) / (\pi\Delta)$ where $\hbar = h/2\pi$, V_F is the Fermi velocity and Δ is the superconducting energy gap.

2.13.12Ginzburg Landau Parameter (k):

It is the ratio of London penetration depth (λ) to coherence length (ξ). $k = \lambda/\xi$. If $k < (1/\sqrt{2})$, the material is type I superconductors, and if $k > (1/\sqrt{2})$, the material is type II superconductors.

2.13.13.Flux Quantization:

The magnetic flux ϕ through a superconducting ring is quantized in terms of h/2e. This phenomenon is said to be flux quantization. A closed superconducting ring (loop) can enclose magnetic flux only in integral multiples of a fundamental quantum of flux. This fundamental quantum of flux is called Fluxon (ϕ_o**)**.

$\phi_o = (h/2e) = [6.626 \times 10^{-34} /(2 \text{ x } 1.602\times10^{-19})] = 2.068\times10^{-15}$ Weber.

The magnetic flux enclosed by a superconducting ring is $\phi = n\, \phi_o = n\, (h/2e)$. The quantization of magnetic flux has been confirmed experimentally. The magnetic flux produced in an ordinary solenoid or transformer is not quantized.

2.13.14Thermal Conductivity:

The thermal conductivity undergoes continuous changes between the normal state and superconducting state of the given superconducting material. It usually lowers in the superconducting phase and approaches zero at very low temperatures of the specimen. This suggests that, superconducting electrons (Cooper pair) possibly play no role in heat transfer.

2.13.15Other properties:

- Periodic Table: T_c is the function of the position of the material in the periodic table.

  Ex: Group V and VII metals are having high T_c and Group VI and VIII metals are having low T_c.

- **Skin Effect:** In the superconducting state, superconductor expels the magnetic flux. Hence, DC current flows only over the outer surface of the superconductor. This effect is called Skin effect.
- **Crystal Structure:** The superconducting property of the material is a function of its crystal structure.

  Ex: White Tin – Tetragonal – superconductor, Grey Tin – Cubic – not a superconductor

  Beryllium – Thin film – superconductor, Beryllium – Thick film – not a superconductor

  Bismuth – Thick film – superconductor, Bismuth – Thin film – not a superconductor
- **Impurities:** The magnetic properties of the superconductors are modifies by the addition of impurities.
- Generally good electrical conductors are not good superconductors Ex: Copper, Gold.Also good superconductors are not good electrical conductors at room temperature Ex: Zinc, Lead.
- Generally, monovalent metals (Cu, Ag, and Au), ferromagnetic metals (Fe, Co, and Ni) and antiferromagnetic metals (CoO, NiO) are not good superconductors.
- When the size of the material is reduced below 10^{-6} m, the superconducting property is modified.
- The value of Tc is unaffected by the frequency.

2.14 THEORY OF SUPERCONDUCTIVITY

A number of theories have been put forward to describe the phenomenon of superconductivity *viz*. London theory, Ginzburg and Landau theory and BCS theory. BCS theory explains most of the phenomena associated with superconductivity in natural manner but alone could not explain the anomalous behavior of HTSC. To explain the HTSC, several theories have been proposed like Resonance Valence Bond (RVB) theory, Interlayer tunneling model, Boson Fermion model *etc*. None of these models provide satisfactory explanation of the anomalous features of HTSC oxides. However, the efforts to understand the HTSC oxides pairing mechanism with enhancing future prospects for new HTSC materials and their novel applications are still under research and are in progress.

The theory of super conductivity can be considered in two stages like macroscopic and microscopic. The first stage is to accept the macroscopic description of a superconductor as a body which has no magnetic and electric fields ($\mathbf{B} = 0 = \mathbf{E}$) and to develop the thermodynamic consequences without making any assumptions. The second stage is the microscopic description of a superconductor formulated by Bardeen, Cooper and Schrieffer (BCS) on the basis of advanced quantum theory.

2.14.1 London's Theory (Macroscopic)

London formulated a more general description of electro-dynamic behavior of superconductor. At an extremely low temperature, vibrations of the nucleus and free electrons of certain atoms slow down so much and they synchronize with the passing waves of electrons in the flow of electric current. When this happens resistance to electric current disappears.

2.14.2The BCS Theory (Microscopic)

BCS theory is a quantum mechanical theory and it was proposed by Bardeen, Cooper and Schrieffer in the year 1957. It comes from the experimental results of superconducting behavior of materials namely (a) Isotope effect and (b) variation of specific heat with temperature.

When we consider the isotope effect, we find that two different isotopes of the same materials exhibit different transition temperatures. Why is the mass of the atom involved in pure electronic phenomenon- superconductivity? Obviously, it means that the motion of the positive ions (phonons) in the lattice has to do something with superconductivity.

Heat capacity measurements provided the first indication of energy gap in superconductors. *i.e.* exponential temperature dependence of specific heat is a hall mark of the system with Δ as an energy gap in the spectrum of allowed energy states separating excited states from the ground states.

BCS theory is based on the interaction of two electrons through the intermediary of phonons. When an electron approaches an ion in the lattice, there is a coulomb attraction between the electron and the lattice ion. This produces a distortion in the lattice and causes an increase in the density of ions in the region of distortion. The higher density of ions in the distorted region attracts, in its turn, another electron. Thus a free electron exerts a small attractive force on another electron through a phonon is called Cooper pair. These pair can drift through the crystal without any scattering over the lattice imperfections without exchanging energy with them, *i.e.* their movement is without resistance.

At normal temperatures, the attractive force is too small and pairing of electron does not take place. Each cooper pair consists of two electrons of opposite spins and momenta. At very low temperatures, such pairing is energetically advantageous. In a typical superconductor, the volume of the given pair encompasses as many as 10^6 other pairs. These dense clouds of Cooper pairs form a collective state where strong correlations arise among the motions of all pairs because of which they drift cooperatively through the material. Thus superconducting state is an ordered state of the conduction electrons. The motion of all Cooper pairs is the same. Either they are at rest, or if the superconductor carries a current, they drift with identical velocity. Since the density of Cooper pairs is quite high, even large currents require only small velocity. The small velocity of Cooper pairs combined with their precise ordering minimizes collision processes. The extremely rare collision of Cooper pairs with the lattice leads to vanishing resistivity.

Electron-Lattice- Electron (Electron-Phonon- Electron) Interaction:

Bardeen found that, an electron moving through a crystal lattice has a self energy accompanied with “Virtual Phonon” of energy hʋ. Let us consider an electron of wave vector $\mathbf{K}_1$ moving through a crystal lattice. This electron can distort the lattice and a virtual phonon of wave vector $\mathbf{q}$ is emitted. Thus $\mathbf{K}_1$ is scattered as ($\mathbf{K}_1 - \mathbf{q}$). Consider a second electron of wave vector $\mathbf{K}_2$ is moving through a crystal lattice and close to the previous electron wave vector $\mathbf{K}_1$. Now the virtual phonon q is absorbed by this second electron and $\mathbf{K}_2$ is scattered as ($\mathbf{K}_2 + \mathbf{q}$). Thus the second electron interacts with the first electron *via* lattice deformation and the exchange of virtual phonon between the two electrons. This type of interaction is called **Electron-Lattice- Electron (Electron-Phonon- Electron) Interaction.**

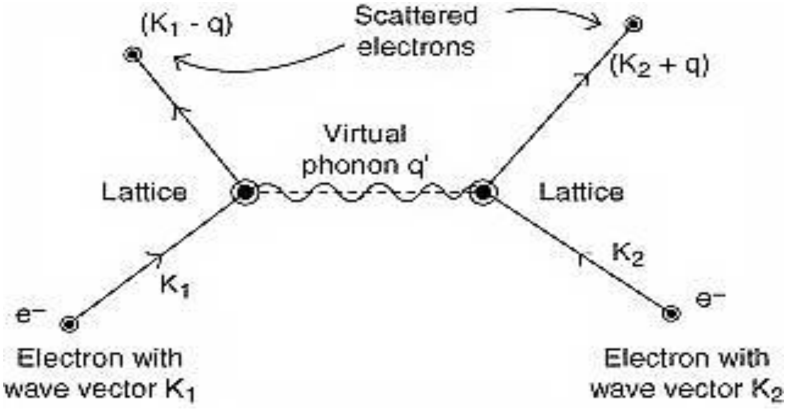


Fig. 2.27 Electron-Phonon- Electron Interaction

Total wave vector before interaction = Total wave vector after interaction

$$\mathbf{K}_1 + \mathbf{K}_2 = (\mathbf{K}_1 - \mathbf{q}) + (\mathbf{K}_2 + \mathbf{q}) = \mathbf{K}_1 + \mathbf{K}_2$$

Thus net wave vector of the two electrons (pair of electrons) is conserved and the momentum is transferred between these two electrons. So these two electrons together form a pair called Cooper pair.

The energy of the Cooper pair (E_{cp}) is less than the total energy of separated electrons (E_{sep}) in the normal state. ($E_{cp}<E_{sep}$). So the binding energy of the Cooper pair (E) is the energy difference between the normal state and paired state ($E=E_{sep}-E_{cp}$). Also this energy is required to break the Cooper pair. At $T > T_c$, all the free electrons are in the separated position. At T $<T_c$, electron-lattice-electron interaction is stronger than electron-electron repulsive coulomb interaction. So the free electrons tend to pair up (formation of Cooper pair). Then the pairing is complete at T = 0 K.

Existence of Energy Gap:

Fermi Energy (E_F) is the highest energy to which the electrons are filled in a metal. The corresponding energy level is called **Fermi Energy level**. The loci of the end points of the wave vectors corresponding to the Fermi level trace out a surface called **Fermi Surface**. The binding energy of the Cooper pair ($E = E_{sep} - E_{cp}$) appears as the superconducting energy gap at the Fermi surface. At the Fermi surface, the normal electron states are above the energy gap while superconducting states are below the energy gap. According to BCS theory, this energy gap at absolute zero is predicted to be $\Delta = 3.5\ K_BT_c$ where K_B is the Boltzmann constant.

The width of the energy gap (Δ) is a function of temperature. At T $>T_c$, Cooper pairing is dissolved and hence the energy gap reduces to zero ($\Delta= 0$). At T = 0 K, all the free electrons are formed in a Cooper pair and hence the energy gap reaches the maximum ($\Delta=$ maximum)s. The energy gaps for typical superconductors ($\sim 10^{-4}$eV) are much smaller than the energy gaps for typical semiconductors (~1eV). This prediction of the superconducting energy gap was verified experimentally by studying absorption of electromagnetic radiation by superconductors.

Major accomplishments of BCS theory

- BCS theory explained Meissner effect, coherence length, penetration depth, flux quantization and energy gap parameter.
- It has solved the problem of electron energy where there is attractive interaction.
- It gives the expression for the transition temperature (McMillan's formula),

$T_c = 1.14\ \theta_D \exp\left[-1/V.N(E_F)\right]$ where θ_D - Debye temperature, V- net attractive potential of the electron, $N(E_F)$ – density of states at the Fermi level.

2.15 TYPES OF SUPERCONDUCTORS

The superconductors are classified into two categories on the basis of magnetic behaviour.

2.15.1Type I (Soft) Superconductor:

In type I (soft) superconductors, the transition from superconducting state to normal state in the presence of a magnetic field occurs sharply at the critical value H_c. They exhibit a complete Meissner effect or perfect diamagnetism. That is, they are perfectly diamagnetic below H_c and completely expel the magnetic field from the interior of the superconducting phase. The negative sign is introduced in the intensity of magnetization value to represent the diamagnetic property of the superconductor. The material produces a repulsive force to the applied magnetic field ($H < H_c$) and the material does not allow the magnetic field to penetrate through it. Hence the material behaves as a superconductor.

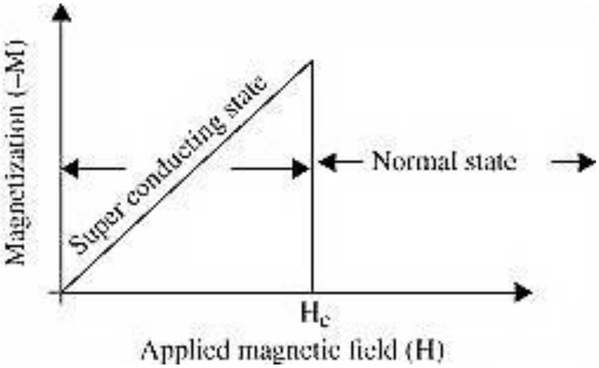


Fig. 2.28 Type I (Soft) Superconductor

At H_c, the repulsive force is zero and hence, the material behaves as a normal conductor. The magnetic field can penetrate only the surface layer and current can only flow in this layer. So type I superconductors are poor carriers of electrical current. The critical field Hc is relatively low for type I superconductors. They would generate field of about 0.01 Wb/m^2 (100 G) to 0.2 Wb/m^2 (2000 G) only. Ex: Lead, Tin, Mercury, Indium, Aluminum.

2.15.2Type II (Hard) Superconductor:

Type II (hard) superconductors are characterized by two critical fields Hc_1 and Hc_2 and the transition from superconducting state to normal state occurs gradually as the applied magnetic field is increased from H_{c1} and H_{c2}. The material produces a repulsive force to the applied magnetic field ($H < H_{c1}$) and the material does not allow the magnetic field to penetrate through it. Hence the material behaves as a superconductor up to lower critical field (H_{c1}). At Hc_1, the repulsive force decreases and magnetic field lines begin penetrating the material. As the applied magnetic field increases further, the magnetic flux lines, through the material, also increases. At upper critical field H_{c2}, the repulsive force is zero and the external magnetic field has completely penetrated.

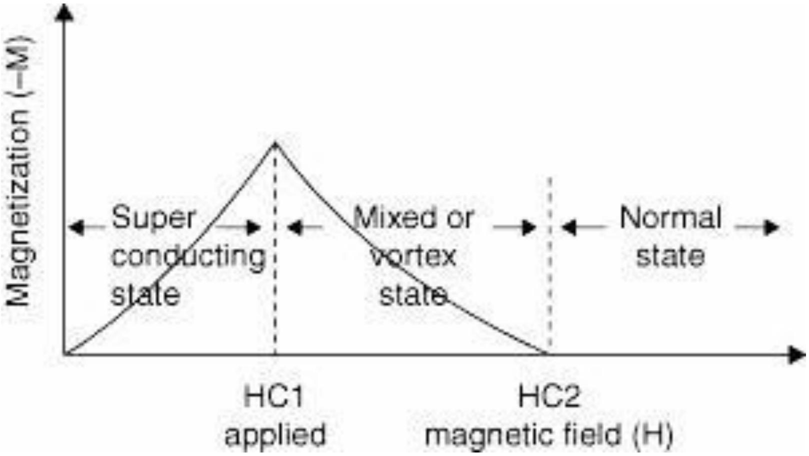


Fig. 2.29Type II (Hard) Superconductor

Hence, the material behaves as a normal conductor. In the region between H_{c1} and H_{c2}, the material is in a magnetically mixed state but electrically it is a superconductor. The value of H_{c2} is about 20 to 50 Wb/m^2, so they can carry larger currents when the magnetic field is between H_{c1} and H_{c2}.

Once the magnetic field is created by superconducting solenoid at $T < T_c$, it does not require electrical power to maintain it. So type II superconductors can carry high supercurrent densities in high magnetic fields which are of great commercial importance. Ex: Niobium (Nb), (Nb_3Sn).

Flux quantization occurs in **type II Superconductors** subjected to a magnetic field. When $H < H_{c1}$ the material behaves as a superconductor. So it can't allow any magnetic field. When $H > H_{c2}$ the material behaves as a normal conductor. So it allow magnetic field completely. When $H_{c1} < H < H_{c2}$ the material is in mixed state i.e. magnetically mixed but electrically superconductor. So it allow magnetic field partially. The bulk material is superconductor but it is threaded by very thin filament of normal state. This thin filament of normal state serves as the path of the magnetic field (magnetic field can penetrate through this path only). This thin filament does not contain any cooper pairs and surrounded by vortices of super current. A flux line together with its current vortex is called **Fluxoid**. Itis a small vertical region of normal phase surrounded by superconducting phase. Small quantum of flux (fluxon) can penetrate through this and super currents circulate around it. The magnetic field through fluxoid is quantized because of the phase properties of the magnetic vector potential in quantum electrodynamics.

2.15.3. Differences between Type I and Type II Superconductors

Type I Superconductor	**Type II Superconductor**
Magnetization drops suddenly, when the specimen is magnetized..	Magnetization drops gradually.
They exhibit complete Meissner effect (diamagnetism).	They do not exhibit complete Meissner effect.
There is only one critical magnetic field (H_c)	There are two critical magnetic fields (H_{c1} and H_{c2})

Type I Superconductor	Type II Superconductor
No mixed state is present.	Mixed state is present.
The value of critical field is low (Hc = 0.1 to 2 T)	The value of critical field is high (H_{c2} = 20 to 50 T)
They are called soft super conductors since they require small magnetic field to destroy the superconductivity.	They are called hard super conductors since they require large magnetic field to destroy the superconductivity.
When the applied magnetic field is reduced below critical field, the material again acquires superconductivity i.e. reversible	Here the process is irreversible
Ginzburg Landau parameter is low: $k < (1/\sqrt{2})$.	Ginzburg Landau parameter is high: $k > (1/\sqrt{2})$
Ex: Lead, Tin Mercury *etc*.	Ex: Niobium and its alloys, transition metals

2.16 JOSEPHSON EFFECT

Josephson junction is the sophisticated sandwich structures of superconducting films usually of Niobium, separated by extremely thin insulating oxide layer of thickness approximately equal to 10 Å. In 1962, B.D.Josephson predicted theoretically that Cooper pair tunneling is possible through a very thin insulating layer that is sandwiched between two superconducting films. The tunneling may be made to result either a dc voltage (D.C Josephson effect) or with an ac voltage (A.C Josephson effect) across the tunnel barrier. This phenomenon is known as Josephson Effect. When a constant nonzero dc potential is maintained across the barrier, ac component exists along the dc component, whereas the dc component persists even after the removal of applied dc potential.

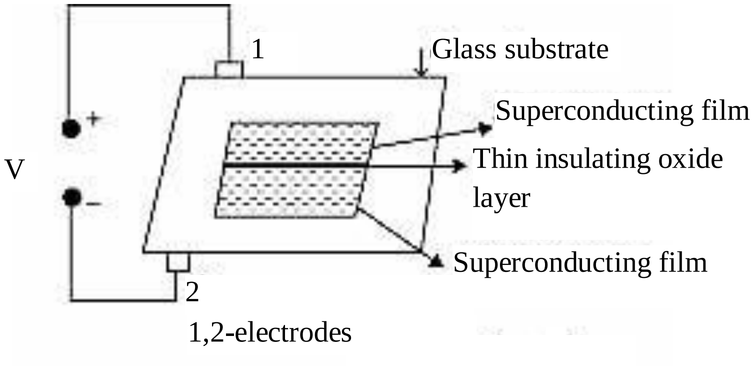


Fig. 2.30 Josephson junction

2.16.1.D.C. Josephson Effect

In a Josephson junction, in the absence of any external electric field across the barrier, there will be tunneling of electrons from one superconducting film of higher electron density to the other superconducting film of lower electron density. This tunneling of electrons results in a dc voltage across the barrier. This effect is called D.C Josephson effect.

Normally, the Cooper pair of electrons present in the same superconducting material is in the same phase. But the Cooper pair of electrons present in two different superconducting materials need not be in the same phase.

Let δ be the phase difference between the Cooper pair of electrons of the two different superconducting materials. Then the dc current density flowing through the junction is given by $J = J_0.\sin\delta$, where J_0 is the maximum current density flowing through the insulator and is a constant. The dc voltage across the barrier is given by, $V = \pi\Delta / 2e$, where Δ – energy gap parameter, 2e – charge on the Cooper pair.

2.16.2. A.C. Josephson Effect

When a constant nonzero potential (V) is maintained across the barrier of the Josephson junction, an alternating super current flows through the barrier in addition to the dc current. This phenomenon is known as ac Josephson Effect. The application of external potential makes the Josephson current to be time dependent. i.e current changes with time, leads to ac current.

This current flow is not due to the tunneling of cooper pairs but it is due the tunneling of single electrons and the junction is now acts like a normal conductor. During the conversion of dc to ac, the cooper pairs are separated into single electrons with the help of an applied dc voltage and the time interval is about 10^{-11} sec. In this short period, there is no current flow across the junction.

The frequency of alternating current is given by $\upsilon = 2eV/ h$ or $\upsilon = K_J V$ where $K_J = (1/ \phi_o) = (2e/h) = 4.835 \times 10^{14}$ Hz.V^{-1} is Josephson Constant (It is the reciprocal of fluxon) and V is the nonzero dc potential. When V = 1 mV; $\upsilon = 4.835 \text{ X } 10^{11}$ Hz which is the frequency of microwaves (This principle is used in the microwave oscillator). By measuring the frequency, the value of e/h and photon energy 2eV can be accurately measured.

2.16.3 I-V Characteristics of Josephson junction

Let V_0 be the applied dc potential and V_c be the minimum dc potential required to produce ac Josephson effect.

(i) When $V_0 = 0$, there is a constant dc current (i_c) flows through the junction. (DC Josephson effect)

(ii) When $V_0 < V_c$, again we have only the constant dc current flows through the junction.

(iii) When $V_0 > V_{c,}$ the junction has finite resistance and the current oscillates with an frequency by $\upsilon = 2eV_0/ h$ (AC Josephson effect)

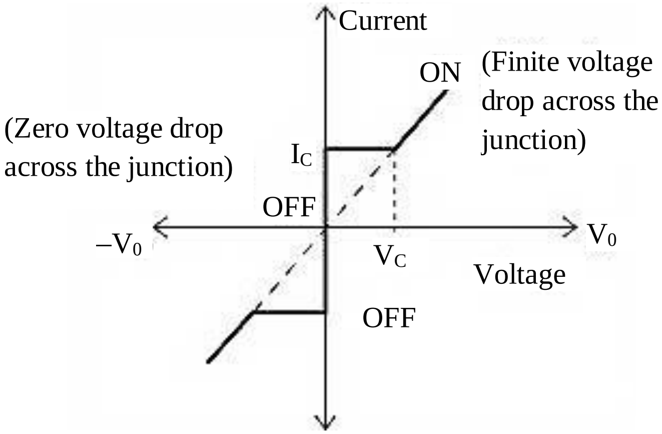


Fig. 2.31I-V characteristics of Josephson junction

2.16.4. Applications of Josephson junction:

1. Josephson junction is capable of switching signals from one circuit to another with a switching time of I picoseconds (10^{-12}sec). (ie) it acts as a fastest switching element.
2. Moreover, it is cable of storing information, smaller in size, lighter in weight; hence they may provide the basis for the architecture of the fastest computers in future. Josephson junction computer has been fabricated with a volume of 1.25 X 10^{-4}m^3 and a cycle time of 2 nano sec; which is 10 times faster than the high speed semiconductor based computer.
3. It is used to produce microwave with a frequency 10^{12} Hz, when the applied potential is 1 mill Volt.
4. It is used in the precision determination of ($\hbar$/e) .
5. The A.C Josephson effect is used to measure very low temperature based on the variation of frequency of the emitted radiation with temperature.
6. Josephson junction is used in the construction of SQUIDs. SQUIDs are used in Magnetocardiography (MCG), Magneto encephalography (MEG) and in magnetic resonance imaging (MRI).

2.17 GENERAL APPLICATIONS OF SUPERCONDUCTORS

2.17.1Cryotron:

It is a magnetically operated current switch where the current through one superconducting element can be altered by the magnetic field produced by another superconducting element. Consider a superconducting material (Tantalum-Ta) is surrounded by another superconducting material (Niobium-Nb) and the whole arrangement is placed in a low temperature bath (T <Tc). The critical field of Tantalum–Ta is less than the critical field of Niobium-Nb.

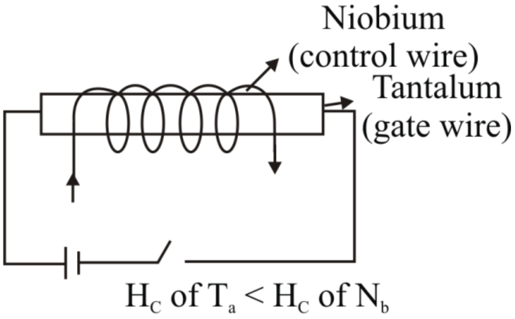


Fig. 2.32 Cryotron

At an operating temperature, the magnetic field produced by the superconducting material Niobium-Nb may exceed the critical field of Tantalum–Ta. Hence, the superconducting material Tantalum–Ta becomes normal conductor since the critical field of Tantalum–Ta is less than the critical field of Niobium-Nb. But Niobium-Nb will not become normal conductor. Hence, the current in material Ta can be controlled by the current in Niobium-Nb and the system can acts as a relay or switching element.

2.17.2Magnetic Levitation (MAGLEV):

The most important application of superconductors is maglev or Magnetic Levitation Train. Maglev coaches do not slide over steel rails but float on a 4 inch air cushion over a strongly magnetized track. The principle behind this is electromagnetic forces between the superconducting magnets. This train consists of superconducting magnets placed on each side of the train. It can run in a guide way (rail) which consists of series of "8" shaped coil. Initially the train slide on the rails. When the train moves, the superconducting magnets on each side of the train will induce a current in the "8" shaped coils kept in the guide way. This induced current generates a magnetic force in the coils in such a way that, the lower half of the "8" shaped coil has the same magnetic pole as that of the superconducting magnet in the train, while the upper half has the opposite magnetic pole. Hence, the total upward magnetic force acts on the train and the train is levitated or raised above the wheels. Now the train floats in the air. By an alternatively changing the poles of the superconducting magnets in the train, alternating currents can be induced in the "8" shaped coils. Thus alternating series of north and south magnetic poles are produced in the coils, which pulls and pushes the superconducting magnets in the train and hence the train is further moved. Because of no mechanical friction, speeds up to 500 Km/hr can be easily achieved.

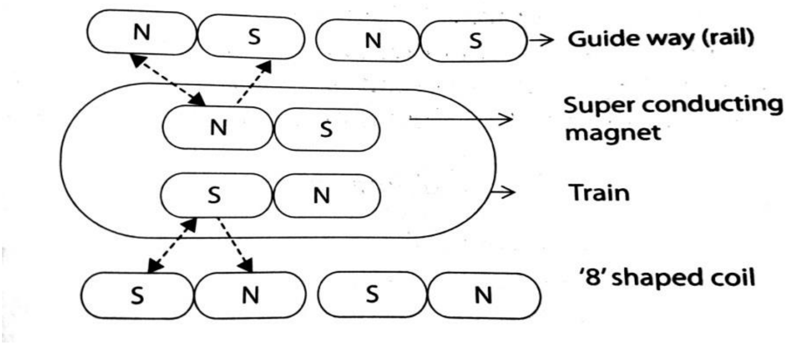


Fig. 2.33Magnetic Levitation Train

2.17.3SQUIDS

SQUIDS means Superconducting Quantum Interference Device and also called Double Junction Quantum Interferometer. It can measure the feeble fields (weak magnetic fields) corresponding to extremely small electric currents generated by heart and brain. The principle behind this is small change in magnetic field produces variation in the flux quantum.

It consists of superconducting ring or loop with “weak link or junction”. Weak link is a thin region which has very low critical current I_c. Weak link acts as a gate, since it allows super current at $I < I_c$ and it blocks super current at $I > I_c$. Weak link is prepared such that it allows only a single fluxon.

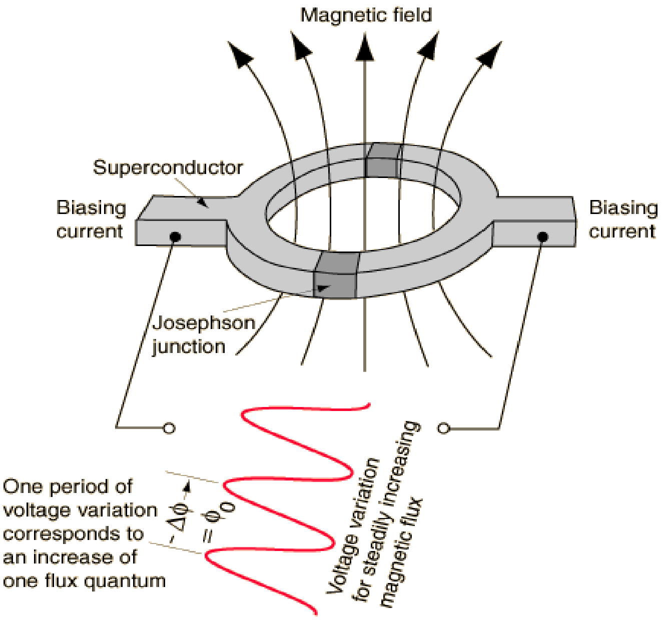


Fig. 2.34Superconducting Quantum Interference Device

Letusconsiderasuperconductingloopwithtwo parallel Josephsonjunctions and no voltage is applied to the junction. Consider high coherent super current coming along the path 1 is divided along two paths ‘a’ and ‘b’ and again merge into one along the path 2.

The super current along the path 1 is represented by a wave function with a phase angle is $I = I_0 \sin\delta_0$

where I_0 is the maximum current and δo is the initial phase angle.

Current along the path a is $I_a = I_0 \sin\delta_a = I_0 \sin[\delta_0 - (\pi\Phi/\Phi_0)]$

Current along the path b is $I_b = I_0 \sin\delta_b = I_0 \sin[\delta_0 + (\pi\Phi/\Phi_0)]$

whereδ_a and δ_b are the phase difference between the paths 1&a and 1&b respectively. Φ is the external applied flux. Φ_0 is fluxon.

In the absence of applied magnetic field (**B** = 0), phase difference is zero ($\delta_b - \delta_a = 0$). In the presence of applied magnetic field (**B** ≠ 0), phase difference is not zero ($\delta_b - \delta_a \neq 0$) ($\delta_b - \delta_a = 2\pi\Phi/\Phi o$)

Total current along the path 2 is $I_T = I_a + I_b$

$$= Io \{\sin[\delta_0 - (\pi\Phi/\Phi_0)] + \text{Sin}[\delta_0 + (\pi\Phi/\Phi_0)]\}$$

$$= 2\, I_0 \sin\delta_0 \cos(\pi\Phi/\Phi_0)$$

This equation represents the interference of the two coherent super currents flowing along the paths 1 and 2. I_T varies with $\cos(\pi\Phi/\Phi_0)$ and the condition for current maxima is $(\pi\Phi/\Phi_0) = \pi, 2\pi, 3\pi, 4\pi$, etc. So current maxima occurs by the magnetic flux increases by a fluxon.

In optical interference, light maxima are produced by phase difference (by path difference) between the coherent lights. But in super current interference, current maxima are produced by phase difference (by applied magnetic field difference) between the coherent super currents.

The graph shows the typical variation of Ic as a function of magnetic field. Each oscillation corresponds to a fluxon. The quantum interference of super currents permits to count fluxon in a direct way.

When an external weak magnetic field produced by heart or brain is applied perpendicular to the plane of the superconducting ring, current is induced at the Josephson junction and produces interference pattern. The induced current flows around this ring, so that, the magnetic flux in the ring can have quantum values of flux, which corresponds to the value of magnetic field applied.

2.17.4Applications of SQUID:

1. By measuring the current through the SQUID, one can measure the value of e/h accurately.
2. SQUID is used to detect very tiny magnetic field and hence it is used as a magnetometer.
3. MRI (Magnetic Resonance Imaging) scanning, employing SQUID, is used to measure the field in the micro Tesla region, where as the ordinary MRI is used to measure the field from one to several Tesla.
4. Based on these sensors we have, magnetocardiography (MCA), Magneto encephalography (MEG), etc.
5. They can be used in the finest precision instruments at the forefront of meteorology.
6. A novel application of SQUID is the magnetic marker monitoring method, which is used to trace the path of orally applied drugs.

2.17.5.Engineering:

(a). When superconductors are used as electric cables, the losses (due to I^2R) are avoided and electrical power transmission can be done at a lower voltage level or a long distance.

(b). Superconductors are used in relay switching circuits since superconducting property can be easily destroyed.

(c). Superconductors are used in gating circuits, storage device in computers etc.

(d). Superconducting coils in transformers and electrical machines much stronger magnetic fields than coils made by other materials since losses due to eddy current and hysteresis are very low. So the size and weight of the motors and generators will be reduced and they will be more efficient.

2.17.6Medical:

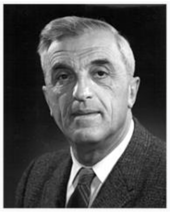
(a). Superconductors are used in NMR (Nuclear Magnetic Resonance) imaging equipments which is used for scanning purposes.

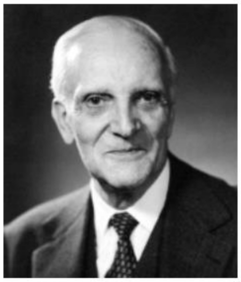
(b). they are used to detect the brain tumor and defective cells.

(c). they are used to separate the damaged cells from the healthy cells.

(d). they are used in Magneto-hydrodynamic power generation to maintain plasma in the body.

Just to know

 **Dr.FELIXBLOCH** (1905-1983) Nobel Prize for Physics (1952)	 **HEIKE KAMMERLINGH ONNES**
His doctoral thesis established the quantum theory of solids, using Bloch waves to describe electrons in periodic lattices. His additional excellent contributions are nuclear induction and nuclear magnetic resonance, which are the underlying principles of MRI and Magnon.	Dutch scientist **Heike KammerlinghOnnes** was awarded the 1913 Nobel Prize in Physics for his work on the properties of matter at low temperatures. He discovered superconductivity



LÉON BRILLOUIN
(1889-1969)

Concept of Brillouin Zone

The concept of a Brillouin zone was first developed by Léon Brillouin (1889-1969), a French physicist. During his work on the propagation of electron waves in a crystal lattice, he introduced the concept of Brillouin zone in 1930. Quantum mechanical perturbations techniques by Brillouin and by Eugene Wigner resulted in what is known as the Brillouin-Wigner formula. He also made contributions to quantum mechanics, radio wave propagation in the atmosphere, solid state physics, and information theory. He also studied the propagation of monochromatic light waves and their interaction with acoustic waves, i.e., scattering of light with a frequency change, which became known as Brillouin scattering.

A Brillouin zone is defined as a Wigner-Seitz primitive cell in the reciprocal lattice. The first Brillouin zone is the smallest volume entirely enclosed by planes that are the perpendicular bisectors of the reciprocal lattice vectors drawn from the origin. The concept of Brillouin zone is particularly important in the consideration of the electronic structure of solids. There are also second, third, *etc*., Brillouin zones, corresponding to a sequence of disjoint regions (all with the same volume) at increasing distances from the origin, but these are used more rarely. As a result, the first Brillouin zone is often called simply the Brillouin zone.

Solved Examples

Problems:

1. Find the relaxation time of conduction electrons in a metal of resistivity 1.54 × 10^{-8} Ω-m, if the metal has 5.8 × 10^{28} conduction electrons/m3.

   Sol: Given data are:

   Resistivity of the metal, $\rho = 1.54 \times 10^{-8}$ Ω–m

   Number of conduction electrons, $n = 5.8 \times 10^{28}/\text{m}^3$

   Relaxation time, τ = ?

$$\sigma = \frac{ne^2\tau}{m} \text{ (or) } \tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho}$$

$$= \frac{9.11 \times 10^{-31}}{5.8\times10^{28} \times [1.602\times10^{-19}]^2 \times 1.54\times10^{-8}}$$

$$= = \frac{9.11\times10^{-31}}{5.8\times(1.602)^2 \times 1.54\times10^{-18}} = \frac{9.11\times10^{-13}}{22.92}$$

$$= \frac{911\times10^{-15}}{22.92} = 39.747\times10^{-15}\text{s}$$

2. The electrical resistivity of copper at 20°C is 2x10^{-8} ohm-metre. Calculate its thermal conductivity

   Sol: Given data are

   Resistivity of the metal, $\rho = 2 \times 10^{-8}$ Ω–m

   Temperature = 20°C = 20 + 273 = 293 K

   Thermal conductivity K ?

   According to theWiedemann Franz law

   $K/\sigma = LT$ where $\sigma = 1/\rho$ and $L = 2.44\ 10^{-8}$ WΩ/K^2

Hence $K = LT/\rho = (2.44\ 10^{-8} * 293) / 2 \times 10^{-8}$

$= 357.46$ W/m/K

3. Lead in the superconducting state has critical temperature 6.2 K at zero magnetic field and a critical field 0.064 T. Calculate the critical field at 4 K.

   Sol: WKT $H_C = H_0 [1 – (T / T_C)^2]$

   Here, we have

   $H_0 = 0.064$ T, $T_c = 6.2$ K and T = 4 K

   $H_c = 0.064 [1 – (4/6.2)^2]$

   $H_c = 0.037$ T

4. The critical values of magnetic field for niobium are 2×10^5 A/m and 10^5 A/m at 0 K and 8 K respectively. Determine its critical temperature.

   Sol: WKT $H_C = H_0 [1 – (T / T_C)^2]$

   Hence $T_C = T / [1 – (H_C / H_0)]^{½}$

   Here, we have

   $H_0 = 2 \times 10^5$ A/m, $H_C = 10^5$ A/m, T = 8 K

   $T_C = 8 / [1- (10^5 / 2 \times 10^5)]^{½}$

   $T_C = 11.312$ K

5. The critical temperature of mercury is 4.153 K for its one isotope of mass 200.59 amu. Calculate the critical temperature of mercury for its another isotope of mass 204 amu.

   Sol. WKT $T_c \propto [1 / M^{½}]$

   $(T_{c1}) / (T_{c2}) = (M_2 / M_1)^{½}$

   Here, we have

   $T_{c1} = 4.153$ K, $M_1 = 200.59$ amu and $M_2 = 204$ amu

   $T_{c2} = (T_{c1}). (M_1 / M_2)^{½}$

   $T_{c2} = 4.153 (200.59/204)^{½}$

   $T_{c2} = 4.118$ K

6. Determine the critical current that can flow through a long thin superconducting wire of aluminium of diameter 10^{-3} m. The critical magnetic field for aluminium is 7.9×10^3 A/m

   Sol.

   WKT $I_C = 2\pi r H_c$

   Here, we have

   $H_C = 7.9 \times 10^3$ A/m, $2r = 10^{-3}$ m

   $I_C = 3.14 \times 10^{-3} \times 7.9 \times 10^3$

   $I_C = 24.806$ A.

7. Calculate the London penetration depth for lead at 5.2 K if the penetration depth at 0 K is 37 nm. The critical temperature of lead is 7.193 K

   Sol.

   WKT $(\lambda_T / \lambda_o) = [1 - (T/Tc)^4]^{-1/2}$

   Hence, $\lambda_T = \lambda o\, [1 - (T/Tc)^4]^{-1/2}$

   Here, we have

   $\lambda o = 37$ nm, $T_C = 7.193$ K and $T = 5.2$ K

   $\lambda_T = 37\, [\, 1 - (\, 5.2/7.193)^4]^{-1/2}$

   $\lambda_T = 43.4$ nm.

8. Calculate the superconducting energy gap at T = 0 K predicted by the BCS theory for cadmium. T_c for cadmium is 0.517 K

   Sol.

   WKT $\Delta = 3.5\, K_B T_c$

   Here, we have

   $T_C = 0.517$ K, $K_B = 1.38 \times 10^{-23}$ J/K

   $\Delta = (3.5 \times 1.38 \times 10^{-23} \times 0.517) / 1.6 \times 10^{-19}$eV

   $\Delta = 1.56 \times 10^{-4}$eV

9. A D.C voltage of 1 μV is applied across a Josephson junction. Calculate the frequency of the Josephson current generated.

Sol.

WKT $\upsilon = 2eV/ h$

Here, we have

$V = 1\ \mu V = 1 \times 10^{-6}\ V$, $e = 1.6 \times 10^{-19}\ C$ and $h = 6.626 \times 10^{-34}\ J/S$

$\upsilon = (2eV) / h = (2 \text{ x } 1.6 \text{ x } 10^{-19} \times 1 \times 10^{-6}) / 6.626 \times 10^{-34}$

$\upsilon =$ 482.9MHz.

10. The superconducting energy gap for a material is 1.5 x 10^{-4} eV. Compute the wavelength of a photon whose energy is just sufficient to break up a Cooper pair in it.

Sol.

WKT $E_g = h\upsilon = hc/ \lambda$ or $\lambda =(hc) / E_g$

Here, we have

$E_g = 1.5 \times 10^{-4}\ eV$, $c = 3 \times 10^{8}\ m/s$ and $h = 6.626 \times 10^{-34}\ J/S$

$\lambda = (hc) / E_g =.(6.626 \times 10^{-34} \times 3 \times 10^{8}) / 1.5 \times 10^{-4}$

$\lambda = 8.28 \times 10^{-3}$ m.

Physics of Solids

Part-A: Questions

1. Differentiate crystalline and non-crystalline solids
2. How the bands are formed in a solid
3. What is allowed band and forbidden band
4. Draw the band formation diagram
5. Differentiate metal, half metal and semi metal
6. Differentiate conductor and insulator
7. Explain types of semiconductor based on energy band theory
8. How Fermi level play a role in p-type and n-type semiconductor
9. Explain Fermi surface
10. Explain Fermi energy
11. Explain Fermi level
12. Explain density of states with respect to Fermi level
13. Explain the effect of external field on Fermi surface
14. Write some merits of free electron theory
15. Write few drawbacks of classical free electron theory
16. How quantum is better than classical free electron theory
17. Write about zone theory
18. State Wiedemann Franz law
19. Define drift velocity
20. Define collision time
21. Define mean free path
22. Define relaxation time
23. Compare classical and quantum free electron theory
24. Differentiate conductor, semiconductor andinsulator.
25. Explain briefly about the types ofsemiconductor.
26. What do you meant by intrinsicsemiconductor?

27. What do you meant by extrinsicsemiconductor?
28. What do you meant by n-typesemiconductor?
29. What do you meant by p-typesemiconductor?
30. What is doping. Mention the use of doping in n-type and p-typesemiconductors.
31. Explain how the acceptor energy level iscreated
32. Explain how the donor energy level iscreated
33. What is Fermilevel?
34. Define superconductivity.
35. What are superconductors?
36. Define transition temperature of the superconductor.
37. Write short note on occurrence of superconductivity.
38. List out some important physical changes in material that occur at T_c.
39. Explain the formation of cooper pair.
40. What is superconducting energy gap?
41. Define coherence length.
42. What do you mean by persistent current?
43. Explain Meissner effect with diagram.
44. Prove that superconductor is a perfect diamagnet.
45. What do you mean by critical magnetic field of a superconductor? How it varies with temperature?
46. What is Silsbee’s rule?
47. Define penetration depth of a superconductor.
48. What is isotope effect?
49. How the entropy of superconductor vary with temperature?
50. Write a note on specific heat capacity of superconductors.
51. Differentiate energy gap of superconductor from semiconductor.
52. What is skin effect?

53. What do you mean by flux quantization?
54. List out the major accomplishments of BCS theory.
55. Differentiate Type I and Type II superconductors.
56. What is Ginzburg Landau Parameter? Mention its significance.
57. What is the effect of pressure on materials?
58. How can you generate microwaves using Josephson device?
59. Explain the IV characteristics of Josephson current?
60. What is a Josephson device? Explain with diagram.
61. Explain DC and AC Josephson effects.
62. List out applications of Josephson's device.
63. What are high temperature superconductors? Give examples.
64. What are high T_c and low T_c superconductors? Differentiate them.
65. What is Magnetic levitation and explain the function of Maglev train.
66. Define fluxon and fluxoid.
67. What is a cryotron and why it is called so?
68. What is SQUIDS and mention its uses.

Part – B: Questions

1. State and explain Block theorem
2. Explain the band structure of conductor, semiconductor, insulator, semimetals and Half metals
3. State and derive Wiedemann Franz law
4. Explain band theory. Differentiate conductor, semiconductor and insulator based on band theory. Also differentiate intrinsic and extrinsic semiconductors based on bandtheory.
5. Define intrinsic semiconductor. Derive an equation of electron and hole concentration of an intrinsicsemiconductor.

6. Define n-type semiconductor. Derive an equation of electron concentration of n- typesemiconductor.
7. Define p-type semiconductor. Derive an equation of hole concentration of p-type semiconductor.
8. Define Band gap. Explain the determination of band gap of asemiconductor.
9. Explain in detail the BCS theory of superconductors.
10. Explain type I and type II superconductors with graphs and write the differences between them.
11. Explain the important properties of superconductors.
12. Write short notes on high T_c superconductors.
13. Write a note on applications of superconductors with reference to cryotron, maglev and SQUIDS.
14. Write short notes on (i) Meissner effect (ii) Silsbee rule (iii) Isotope effect (iv) Critical magnetic field and (v) Specific heat capacity of superconductors.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF PHYSICS

UNIT – III - PHYSICS FOR ENGINEERS-SPHA1101

UNIT 3

MAGNETISM, LASER FUNDAMENTALS AND OPTO ELECTRONICS

3.0INTRODUCTION

Magnetism was observed as early as 800 B.C is naturally occurring material called load stone. In the modern concept, all materials, viz., metals, semiconductors and insulators are said to exhibit magnetism of different nature. The magnetic properties of solids originate in the motion of electrons and in the permanent magnetic moments of the atoms and electrons. '**Any material that can be magnetised by application of external magnetic field is called a magnetic material**'. Of many types of magnetic materials diamagnetic, paramagnetic, ferromagnetic, antiferromagnetic and ferrimagnetic are the most important from the point of view of practical (Engineering) applications.

BASIC DEFINITIONS

We will first consider the terms and definitions used in magnetism.

i) Magnetic lines of force

These are the lines which are due to the magnetizing field and exist in air even when there is no magnetic substance in the field. It is shown in fig 3.1.

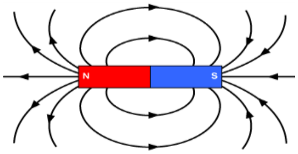


Fig.3.1 Magnetic lines of force

ii) Lines of magnetisation

If a magnetic substance is kept in a magnetic field, it is magnetised by induction and these lines of force passing within the substance due to magnetisation are known as lines ofmagnetisation as shown in fig.3.2.

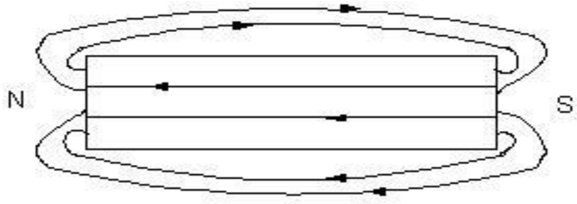


Fig.3.2 Lines of magnetisation

iii) Magnetic Induction or magnetic flux density (B)

In any material is number of lines of magnetic force passing through unit area perpendicularly (μ/A). Its unit is wbm^{-2} or Tesla.

iv)Magnetic field intensity (H)

At any point in the magnetic field is the force experienced by an unit north pole placed at that point. Its unit is Am^{-1}.

v) Magnetic Susceptibility (χ_m)

It is defined as the ratio between intensity of magnetization (I) and the magnetic field intensity (H).

$$\chi_m = \frac{I}{H}$$

vi) Intensity of Magnetization (I)

It is defined as the magnetic moment per unit volume.

Its unit is Am^{-1}.

$$I = \frac{M}{V}$$

vii) Magnetic permeability (μ)

It is defined as the ratio between the magnetic flux density (B) and the magnetic field intensity (H). It is the measure of degree at which the lines of force can penetrate through the material .Its unit is $\mathrm{Hm^{-1}}$

$$\mu = \mu_o \mu_r = \frac{B}{H}$$

viii) Relative Permeability (μ_r)

It is the ratio between the permeability of the medium to the permeability of free space.

$$\mu_r = \frac{\mu}{\mu_o}$$

ix) Relation between (μ_r and χ_m)

When a magnetic material is placed in a magnetic field (H), then two types of lines of induction passes through the material viz

i) Due to magnetizing field (H)

ii) Due to material itself being magnetized by induction (I)

$\therefore$ Total flux density, $B = \mu_o(H + I)$ (1)

We know, $\mu = \frac{B}{H} \Longrightarrow B = \mu H$ (2)

Equating eqn (1) and (2) We get

$$\mu H = \mu_0 (H + I)$$

Since $\mu = \mu_0 \mu_r$ we have

$$\mu_0 \mu_r H = \mu_0 H \left[1 + \frac{I}{H}\right]$$

Here, $\frac{I}{H} = \chi_m$

$$\mu_r = \left[1 + \frac{I}{H}\right] = 1 + x_m$$

$$\frac{I}{H} = x_m$$

x) Force between two poles

Let us consider two poles of pole strength m_1 and m_2 placed at a distance 'r' apart as shown in fig.3.3.

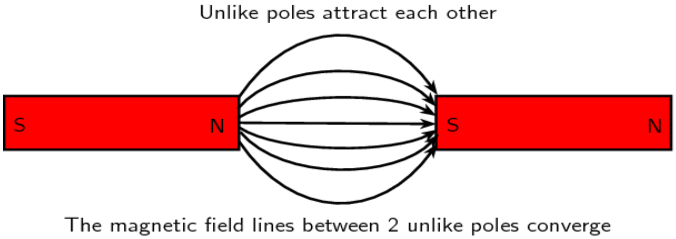


Fig.3.3 Force between two poles

From the coulomb's law of forces, the force of attraction (or) repulsion between the isolated point magnetic poles is proportional to the product of pole strengths and is inversely proportional to the square of the distance between them

i.e, $F \infty \frac{m_1 m_2}{r^2}$

$$F = K \frac{m_1 m_2}{r^2}$$

In terms of unit vector $\hat{r}$ directed from m_1 to m_2

$$F = k\frac{m_1 m_2}{r^2}\hat{r}$$

$$\vec{F} = K\frac{m_1 m_2}{r^3}\vec{r} \qquad \hat{r} = \frac{\vec{r}}{r}$$

Here K is a constant of proportionality

$$K = \frac{1}{4\pi\mu_0\mu_r}$$

Where $\mu_0 \rightarrow$ Permeability in free space $\left(4\pi\times10^{-7}\,H/M\right)$

$\mu_r \rightarrow$ Relative permeability

$$\vec{F} = \frac{m_1 m_2}{4\pi\mu_0\mu_r r^3}\vec{r} \text{ for air } \mu_r = 1$$

$$\therefore \vec{F} = \frac{m_1 m_2}{4\pi\mu_0 r^3}\vec{r} \text{ Newton}$$

3.1Origin of Magnetic Moment – Bohr Magneton

We shall now discuss now what contribute to the permanent magnetic dipole moment of the atomic constituents of matter. Whenever a charged particle has an angular momentum, it contributes to the permanent dipole moment. In general, these are three contributions to the angular momentum of an atom.

i) Orbital angular momentum of the electrons: This corresponds to an permanent orbital angular magnetic dipole moments.

ii) Electron spin angular momentum. This corresponds to electron spin angular momentum.

iii) Nuclear spin angular momentum. This corresponds to nuclear magnetic moments.

i) Orbital angular magnetic dipole moment

Let us consider an electron describing a circular orbit of radius 'r' with a stationary nucleus at the centre as shown in fig.3.4.Let the electron rotate with a constant angular velocity of w_0 radians per second. Electron revolving in any orbit may be considered as current carrying circular will producing magnetic field perpendicular to its plane. Thus the electronic orbits are associated with a magnetic moment. The orbital magnetic moment of an electron in an atom can be expressed interms of atomic unit of magnetic moment called **Bohr magneton**, defined as1 Bohr magneton $= \frac{e}{2m}\frac{h}{2\pi}$

$$= \frac{eh}{4\pi m} = \mu B$$

$$= 9.27 \times 10^{-24} \text{ Ampere } m^2$$

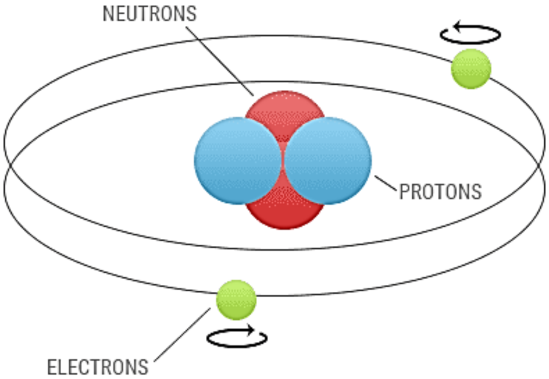


Fig.3.4 The electron describing the circular orbit around the nucleus

ii) Electrons spin magnetic moment

The concept of the electron having an angular momentum has been introduced in order to explain the details of atomic spectra. This angular momentum of the electron is referred to as the spin of the electron. Since the electron has a charge, its spin produces a magnetic dipole moment. According to quantum theory, the spin angular momentum along a given direction is either $+\frac{h}{2\pi} or -\frac{h}{2\pi}$.

Hence the spin dipole moment components along an external field are

$$+\frac{e}{m}\frac{h}{4\pi} = +1 \text{ Bohr Magneton}$$

or

$$-\frac{e}{m}\frac{h}{4\pi} = -1 \text{ Bohr Magneton.}$$

iii) Nuclear magnetic moment

The angular momentum associated with the nuclear spin is also measured in units of $\frac{h}{2\pi}$. The mass of the nucleus is larger that of an electron by a factor of the order of 10^3. Hence nuclear spin magnetic moment is of the order of 10^{-3} Bohr magnetons.For all practical purposes, we assume that the permanent magnetic dipoles arise due to the electron spin ignoring the orbital magnetic moments and the nuclear magnetic moments as their magnitude are small.

Classification of magnetic materials

Now we are going to study the various types of magnetic materials in terms of the magnetic properties of the atomic dipoles and the interactions between them.Very first distinction is based on whether the atoms carry permanent magnetic dipoles or not. Materials which lack permanent dipoles are called diamagnetic. If the atoms of the material carry permanent magnetic dipoles, such a material may be paramagnetic, **ferromagnetic**, antiferromagnetic or ferrimagnetic, depending on the interaction between the individual dipoles. Before studying the properties of different magnetic materials let us classify them based on their properties.

Diamagnetism

An electron moving around the nucleus results in magnetic moment. Due to different orientations of various orbits of an atom, the net magnetic moment is zero in diamagnetic materials. When an external magnetic field is applied the motion of electrons in their orbits changes resulting is induced magnetic moment in a direction **opposite** to the direction of applied field. Thus diamagnetism is a property of all atoms because of the influence of an applied magnetic field on the motion of electrons in their orbits. It is a weak effect and in solids it is often masked by other kinds of magnetism.

Properties of diamagnetic materials

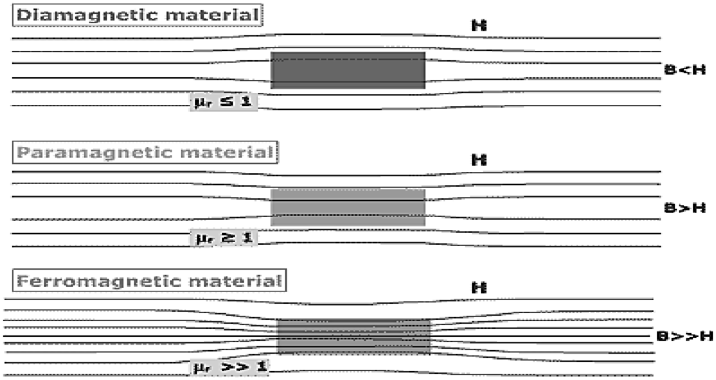


Fig.3.5 The behavior of magnetic materials in thePresence of magnetic field

- Permanent dipoles are absent.
- Effect is weak and often marked by other kinds of magnetism.

- When placed inside a magnetic field, magnetic lines of forces are repelled (Meissner's Effect) as shown in fig .3.5.
- Diamagnetic susceptibility is negative
- Magnetic susceptibility is independent of applied magnetic field strength.
- Relative permeability is slightly less than unity.

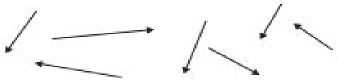
Example: Gold, Germanium, Silicon, etc

Paramagnetism

Each electron in an orbit has an orbital magnetic moment and a spin magnetic moment even in the absence of an external magnetic field. In the absence of external magnetic field the net moments of the atom are arranged in random directions because of thermal fluctuations. Hence there is no magnetization. When external magnetic field is applied, there is a tendency for the dipoles to align with the field giving rise to an induced positive dipole moment. This induced dipole moment is proportional to the field. The induced magnetism is the source of paramagnetism.

Properties of Paramagnetic materials

- Paramagnetic materials posses permanent magnetic dipoles.
- In the absence of an external applied field, the dipoles are randomly oriented. Hence the not magnetization in any given direction is zero.
- When placed inside be magnetic field, it attracts the magnetic lines of force as shown in fig 3.5.
- Paramagnetic susceptibility is positive and depends greatly on temperature as detailed below.
- Spin alignment is random



- Paramagnetic susceptibility is independent of the applied magnetic field strength.

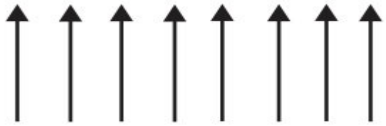
Example: Manganoussulphate, Ferric oxide, Ferrous sulphate, Nickel sulphate etc.,

Ferromagnetism

We have known that the permanent magnetic moment is mainly due to the spin magnetic moment. The net spin magnetic moment of ferromagnetic atoms is of the same order as magnetic moment of paramagnetic atoms. However there exists a large magnetization in ferromagnetic substance even in the absence of external applied field. This is due to **spontaneous magnetization**. There is a special form of interaction called exchange coupling occurring between adjacent atoms.

Properties of ferromagnetic materials

- Due to a special form of interaction called exchange coupling between adjacent atoms, even in the absence of external applied field, exhibits strong magnetization
- When placed inside a magnetic field, it attracts the magnetic lines of forces very strongly as shown in fig 3.5
- Each ferromagnetic material has a characteristic temperature called the ferromagnetic **Curie temperatureT_c.** Ferromagnetic susceptibility depends greatly on temperature above of it, its properties are quite different as shown below.
- Spin alignment is parallel in the same direction



- Exhibits hysteresis
- Consists of a number of small regions which are spontaneously magnetized called domains.

Example: Iron, Nickel, Cobalt etc.,

Antiferromagnetism

In ferromagnetism, we have seen that the tendency for parallel alignment of the electron spins is due to quantum mechanical exchange forces. In certain materials when the distance between the interacting atoms is small, the exchange force produces a tendency for anti parallel alignment of electron spins of neighboring atoms. This kind of interaction is encountered in antiferromagnetic and in ferromagnetic materials. The most characteristic feature of an antiferromagnetic material is the occurrence of a rather sharp maximum in the susceptibility versus temperature curve as shown in fig 3.6 for MnF_2.

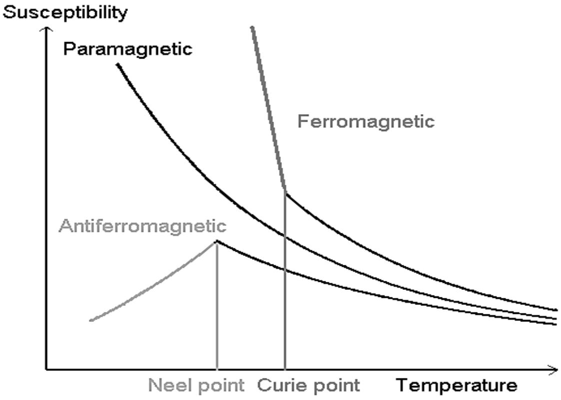


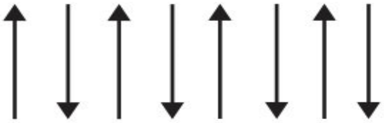
Fig.3.6 The susceptibility of MnF_2

The temperature at which the maximum occurs is called the **Neel temperature,T_N**. Above Neel temperature, the susceptibility is observed to

$x = \frac{C}{T+\theta}$, where C is the curie constant and θ the paramagnetic Curie temperature

Properties of antiferromagnetic materials

- Electron spin of neighboring atoms are aligned antiparallel.
- Antiferromagnatic susceptibility depends greatly on temperature
- Initially susceptibility increases slightly with temperature and beyond Neel temperature the susceptibility decreases with the temperature



- Spin alignment is antiparallel.

Example: MnO, MnS, NiCr etc.,

Ferrimagnetism

This is a special case of antiferromagnetism. The net magnetization of magnetic sublattices is not zero since antiparallel moments are of different magnitudes. Hence ferromagnetic material possesses a net magnetic moment. This moment disappears above a curie temperature T_C analogous to the Neel temperature. Above T_C analogous to the Neel temperature. Above T_C thermal energy randomizes the individual magnetic moments and the material becomes paramagnetic. Fig.3.7 illustrates the variation of susceptibility of with temperature in different magnetic materials.

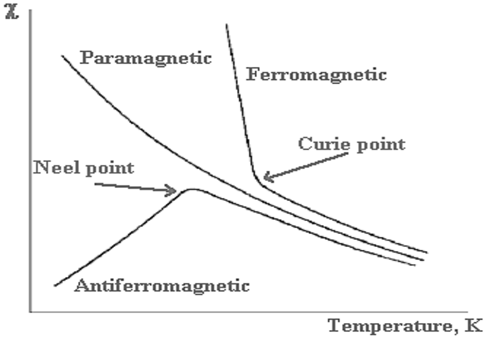


Fig.3.7 Variation of susceptibility of with temperature

Properties of ferrimagnetic materials

- Ferrimagnetic materials possess net magnetic moment.
- Above Curie temperature it becomes paramagnetic while below it behaves as ferromagnetic materials.
- Ferrimagnetic domains become magnetic bubbles to act as memory elements.
- Spin alignment in antiparallel of different magnitudes.

3.2 MAGNETIC HYSTERESIS LOOP

The **Magnetic Hysteresis** loop below shows the behaviour of a ferromagnetic core graphically as the relationship between B and H is non-linear. Starting with an unmagnetised core both B and H will be at zero, point 0 on the magnetisation curve.

- If the magnetisation current, i is increased in a positive direction to some value the magnetic field strength H increases linearly with i and the flux density B will also increase as shown by the curve from point o to point **a** as it heads towards saturation.

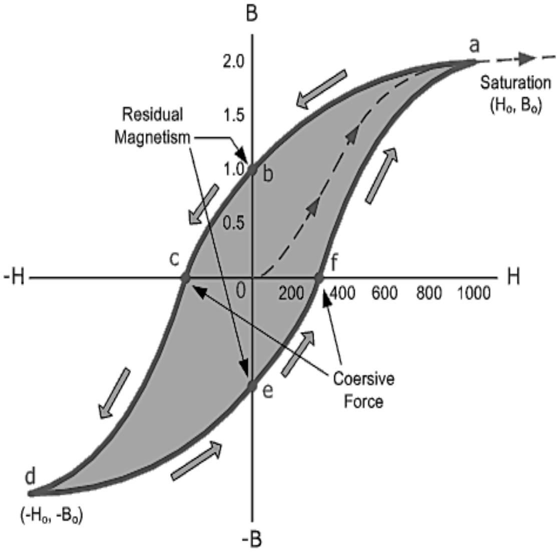


Fig.3.8 Hysteresis loop

- Now if the magnetising current in the coil is reduced to zero, the magnetic field circulating around the core also reduces to zero. However, the coils magnetic flux will not reach zero due to the residual magnetism present within the core and this is shown on the curve from point **a** to point **b**.
- To reduce the flux density at point b to zero we need to reverse the current flowing through the coil. The magnetising force which must be applied to null the residual flux density is called a "Coercive Force". This coercive force reverses the magnetic field re-arranging the molecular magnets until the core becomes unmagnetised at point **c**.
- An increase in this reverse current causes the core to be magnetised in the opposite direction and increasing this magnetisation current further will cause the core to reach its saturation point but in the opposite direction, point **d** on the curve.

- This point is symmetrical to point **b**. If the magnetising current is reduced again to zero the residual magnetism present in the core will be equal to the previous value but in reverse at point **e**.
- Again reversing the magnetising current flowing through the coil this time into a positive direction will cause the magnetic flux to reach zero, point **f** on the curve and as before increasing the magnetisation current further in a positive direction will cause the core to reach saturation at point **a**.
- Then the B-H curve follows the path of **a-b-c-d-e-f-a** as the magnetising current flowing through the coil alternates between a positive and negative value such as the cycle of an AC voltage. This path is called a **Magnetic Hysteresis Loop**as shown in fig.3.8.
- The effect of magnetic hysteresis shows that the magnetisation process of a ferromagnetic core and therefore the flux density depends on which part of the curve the ferromagnetic core is magnetised on as this depends upon the circuits past history giving the core a form of “memory”. Then ferromagnetic materials have memory because they remain magnetised after the external magnetic field has been removed.

However, soft ferromagnetic materials such as iron or silicon steel have very narrow magnetic hysteresis loops resulting in very small amounts of residual magnetism making them ideal for use in relays, solenoids and transformers as they can be easily magnetised and demagnetised.

Since a coercive force must be applied to overcome this residual magnetism, work must be done in closing the hysteresis loop with the energy being used being dissipated as heat in the magnetic material. This heat is known as **hysteresis loss**, the amount of loss depends on the material’s value of coercive force. The area of the loop represents energy loss per cycle per unit volume of the specimen.The product of retentivity (Br) and coercivity (Hc) is known as **Energy product**. It represents the maximum amount of energy stored in the specimen. Therefore for **permanent**

magnets the value of energy product should be very high as shown in fig.3.8. (Magnetic field is produced by objects known as magnets. When these properties of magnetism are not lost throughout time it's known as a**permanent magnet.)**

By adding additive's to the iron metal such as silicon, materials with a very small coercive force can be made that have a very narrow hysteresis loop. Materials with narrow hysteresis loops are easily magnetised and demagnetised and known as soft magnetic materials.

Retentivity

- The value of magnetic field is decreased, but the magnetic flux density (or) magnetic induction (B) does not decrease at the same rate at which it is increased. When $H = 0$, $B \neq 0$ the magnetic induction has a definite value represented by Ob, and is known as **Retentivity**.
- The reason for this that some of the tiny molecular magnets do not return to a completely random pattern and still point in the direction of the original magnetising field giving them a sort of "memory". Some ferromagnetic materials have a high retentivity (magnetically hard) making them excellent for producing permanent magnets.

Coercivity

The applied magnetic field H is reversed and increased gradually till the point C is reached. The magnetic flex density (or) magnetic induction (B) becomes zero at the point C and it is known as **Coercivity.**

Types of magnetic materials

Magnetic materials are classified into two types

(i) Soft magnetic materials

(ii) Hard magnetic materials

i) Soft Magnetic Materials

Materials which are easy to magnetise and demagnetise are called soft magnetic materials. These magnetic materials do not retain the alignment of magnetic domains after the removal of the external magnetic field. The hysteresis loop for soft magnetic material is shown in fig 3.10.

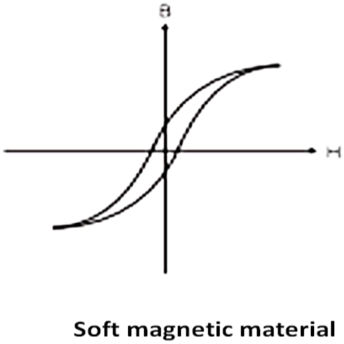


Fig.3.10 Hysteresis loop for soft magnetic material

Properties of soft magnetic materials

- They can be easily magnetized and demagnetized.
- They have high permeability.
- They have low residual magnetism.
- They have low coercive force.
- They have low hysteresis loss.
- The eddy current loss is small due to its high resistivity.

ii) Hard Magnetic Materials

Materials which retain their magnetism and are difficult to magnetise and demagnetise are called hard magnetic materials. These magnetic materials retail permanently the alignment of the magnetic domains after the removal of the external magnetic field. The hysteresis loop for hard magnetic is shown in fig 3.11.

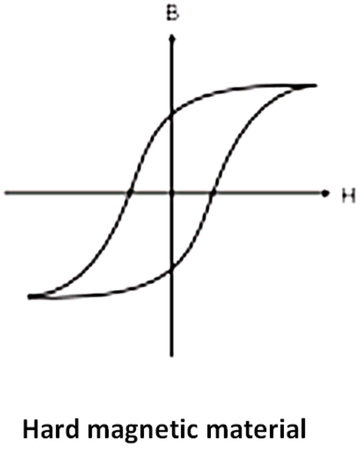


Fig.3.11 Hysteresis loop for hard magnetic material

Properties of Hard Magnetic Materials

- They have low permeability and strong magnetic reluctance.
- They have high retentivity and coercivity.
- They have large hysteresis loop area and large energy losses.
- They require high magnetising force to attain magnetic saturation.
- They possess high value of B-H product.

3.3 DOMAIN THEORY OF FERROMAGNETISM: (WEISS THEORY)

Weiss proposed the concept of domains in order to explain the properties of ferromagnetic materials and its magnetic domains. The group of atoms organised into tiny bounded regions in the ferromagnetic materials are called magnetic domains.

Thus the domain is a region of the ferromagnetic material in which all the magnetic moments are aligned to produce a net magnetic moment in only one direction. Hence behaves likes a magnet with its own magnetic moment and axis. In a demagnetized ferromagnetic material, the domains are randomly oriented.The net magnetisation, in the absence of the field is zero. These domains are separated from other domains by a wall known as the **domain wall or Bloch Wall**.

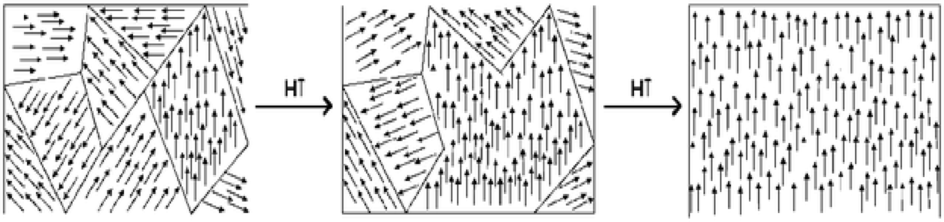


Fig.3.12 Rotation of orientation and increase in size of magnetic domains in response to an externally applied field.

When a magnetic field is applied to this material, the domains that are parallel to the applied field increase in size at the expense of the other domains .The domain size can also grow due to the movement of domain walls. Also, all the other domains aline themselves with the field. This results large net magnetisation of the material as shown in fig.3.12

When the substance is placed in an external field ,magnetism of substance can increase in two different ways

- **By the motion of domain Walls:** The volume of the domains that are favorably oriented with respect to the external field are increased than that are unfavorably oriented as shown in fig 3.13(b). fig 3.13(a) shows domain arrangement in a virgin specimen when no magnetic field is applied.

- **By rotation of domains:** As shown in fig 3.13(c), when the applied external magnetic field is strong, rotation of the direction of magnetization occurs in the direction of the field.

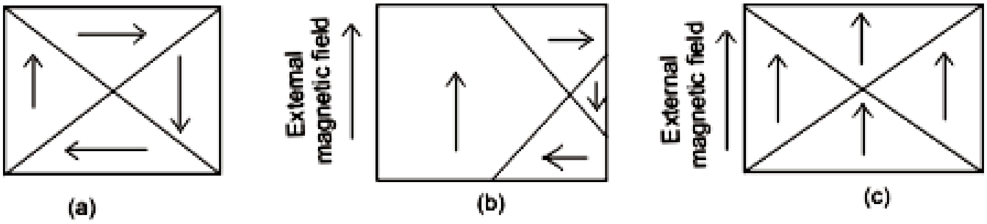


Fig 3.13(a) Random domain alignment(b) Domain wall movement (c) Domain rotation

Hysteresis on the basis of Domains

We know that when the ferromagnetic material is subjected to external field, there is an increase in the value of the resultant magnetic moment due to two processes, viz.

- The movement of domain walls
- Rotation of domain walls

When a small external field is applied, the domain walls are displaced slightly in the easy direction of magnetization. This gives rise to small magnetisation corresponding to the initial portion of the hysteresis curve (OA) as shown in fig.3.14.

Now, if the applied field is removed, then the domains return to its original state, and is known as reversible domains. When the field is increased, large number of domains contributes to the magnetisation and thus the magnetisation increases rapidly with H.

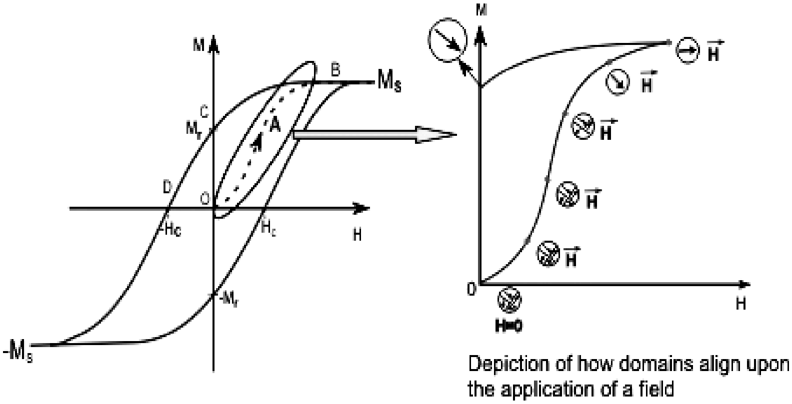


Fig.3.14 Hysteresis on the basis of Domains

Now, even when the field is removed, due to the displacement of domain wall to a very large distance, the domain wall boundaries do not come back to their original position. This process is indicated as AB in fig 3.14 and these domains are called irreversible domains. At point B all the domains have been magnetized along the easy direction. Now when the field is further increased, the domains start rotating along the field direction and the anisotropy energy is stored in the hard direction, represented as BC in the fig. 3.14. Thus the specimen is said to attain the maximum magnetisation. At this position, even after the removal of external field the material possess maximum magnetisation, called residual magnetism (or) retentivity, represented by OD in fig. 3.14

Actually after the removal of the external field, the specimen will try to attain the original configuration by the movement of Block wall. But this movement is stopped due to the presence of impurities, lattice imperfections etc., Therefore to overcome this, a large amount of reverse magnetic field is applied to the specimen, the amount of energy spent to reduce the magnetisation to zero is called as coercivity represented by OE in fig .3.14.

Domain Theory – Experimental verification

The domain concept was experimentally verified by the magnetic powder pattern technique by Bitter In this technique, the colloidal solution of magnetic material is sprayed on the surface of the ferromagnetic material under study as show in fig. 3.15.We know that the strong magnetic field exists near the boundaries.

The colloidal particles are attracted towards well defined lines and show the boundaries between domains as shown in fig. 3.15. Further, these domains move depending on the applied external field. A microscope with high resolving power helps to observe the domain walls and their movements. In this way Bitter confirmed the existence of domains in a magnetic material.

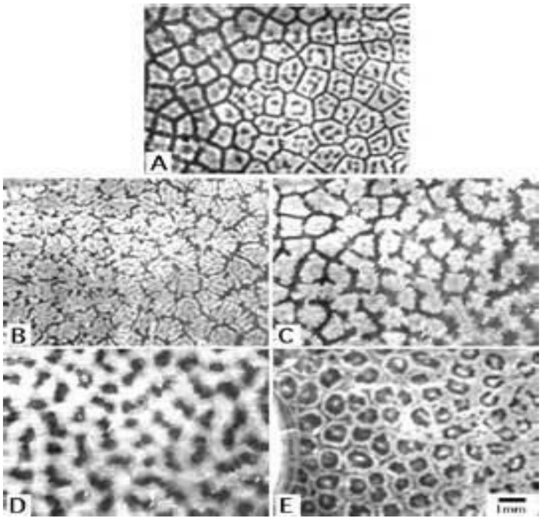


Fig .3.15 Bitter powder pattern

During the growth of domain, the internal energy is due to the net contribution of magnetostatic, anisotropy, domain wall and magnetostriction energies. Let us discuss the above energies and their contribution to internal energy briefly in the following section.

Types of energy involved in the process of domain growth

1. Magnetostatic energy
2. Crystal anisotropy energy
3. Magnetostrictive energy.
4. Domain wall energy.

To study the domain structure clearly, we must know four types of energy involved in the process of domain growth.

1. Magnetostatic Energy:

Magnetostatic energy or magnetic potential energy is the energy present in any ferromaganetic materials when that material produces an external field. This is due to the present of resultant dipole moment in that material even in the absence of external magnetic field. The magnetic energy can be reduced by dividing the specimen into two domains as shown in fig 3.16 (b).

Further subdivisions, as shown in fig 3.16, to N domains, reduces the magnetic energy to 1/N of the magnetic energy of the configuration of fig 3.16 (a).A domain structure shown in fig 3.16 c. has 2000 times of magnetic energies due to the introduction of triangular domain at the top and bottom of the crystal. These triangular domains are called **closure domains.**

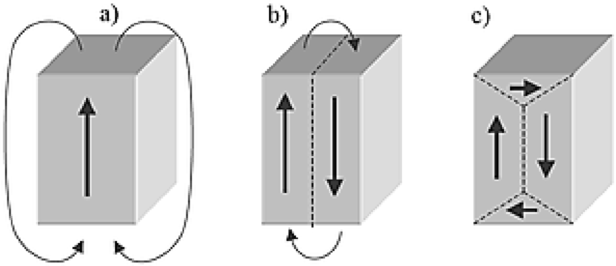


Fig. 3.16 Domain of magnetic materials.

2. Anisotropy Energy:

Energy of megnetisation is the functions of crystal orientation in crystals have easy and hard direction of magnetisation. For e.g. in Bcc iron, the easy direction is [100], the medium direction is [110], and the hard direction is [111].

The excess energy required to magnetise a specimen in a particular direction over that it required to magnetise it along the easy direction is called the crystalline anisotropy energy.

Material	Easy direction	Medium Direction	Hard Direction
Iron	[100]	[110]	[111]
Nickel	[111]	[110]	[100]

Bloch wall or Domain wall energy = Exchange energy +

Crystal anisotropy energy.

3. Magnetostrictive Energy:

The change in the dimension of a ferromagnetic material when it is magnetized is known as magnetostriction. This deformation is different along different crystal directions. So if the domains are magnetized in different directions, they will either expand or shrink. This means that work must be done against the elastic restoring force.

$$\lambda = \frac{\Delta l}{l}$$

Where λ is called the magnetostriction constant

The work done by the magnetic field against these elastic restoring force is called magneto-elastic energy (or) magnetostrictive energy. [It is the energy due to the mechanical stress generated by domain rotations].

- It is used in generation of ultrasonic waves.
- It is responsible for the noise produced by the transformer.
- It depending on the nature of the material the dimensions may either increase or decrease. For Ni rod the length decreases while for a Permalloy the length increases in the presence of magnetic field.

The magnetostriction energy is the energy due to the mechanical stress generated by domain rotation.

4. Bloch or Domain Wall Energy:

Consider two domains in magnetic materials. These two domains are separated by Bloch or domain wall. The domains are opposite in direction. The second domain is obtained by rotating the first domain through 180°. The rotation of the domain is carried out gradually due to the existence of exchange force and anisotropy energy.

The role of exchange force to rotate the dipole which is existing between the adjustment atomic spin is very less.

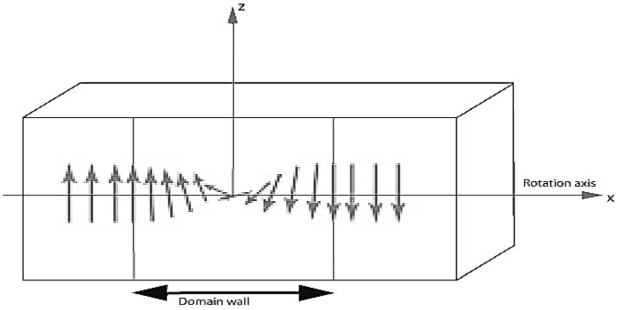


Fig.3.17 Rotation of domains

The gradual rotation of domains in an iron crystal is shown in fig. 3.17.The energy required to rotate domain in its easy direction is less when compared to its hard direction. In order to rotate the domain through an angle of 180°, anisotropy energy required a thickness of domain wall nearly 1A°, where as the exchange energy requires a larger domain wall thickness. However, a minimum wall thickness is required with minimum potential energy at equilibrium thickness condition. Therefore, **the minimum potential energy of domain wall is known as the domain wall energy**. Generally, the thickness of domain is in the order of hundreds of A°. For example, the domain wall thickness of iron is equal to 0-1 µm.

3.4 MAGNETIC BUBBLE MEMORY

The garnet (semi-precious gem of deep transparent redthin filmis growth on a non-magnetic substrate. It is in the form of thin films, in which the magnetic domain get concentrated into circular areas and form circular domains in the order of few microns as shown in fig 3.18. Such circular domains are called magnetic bubbles. This is used in computer memory, using soft magnetic material. They perform binary functions is Computers including logic, memory, counting and switching.

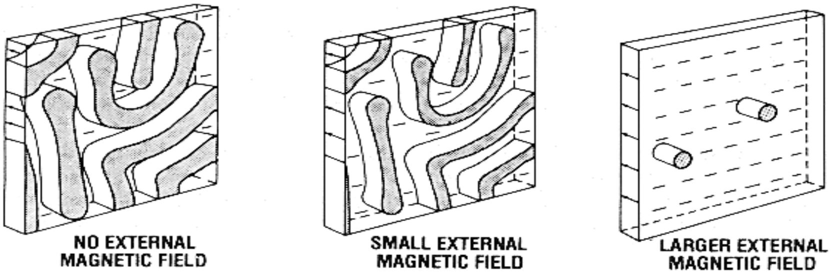


Figure 3.18 Magnetic bubble domains

Bubble Formation:

The garnet thin film is grown in the non-magnetic substrate. The magnetic domains in the garnet film arise from the preferred axis of magnetization perpendicular to the film surface. The magnetic domains are in the form of strips, which are arranged in the ways, like pointing up and down ward. When an external field is applied in a perpendicular direction of the thin garnet film, the magnetic domains (strips) shrink down into tiny circle, known as tiny magnetic domains. The tiny magnetic domains are very small in size when compared to other magnetic storage devices like taps and hard disks. These tiny magnetic domains are known as magnetic bubbles.

The bubble memory was demonstrated in

- Orthoferrites ($RFeO_3$)
- Hexagonal ferrites ($PbFe_{12}O_{19}$)
- Gd-Co and
- Ge-Fe alloy films.

These bubbles are free to move through the film. The formation of bubbles can be observed with a polarizing microscope. The magnetic bubbles can occur only if the magnetic material has a uniaxial anisotropy with the easy axis of magnetisation perpendicular to the film surface. The diameter of this cylindrical magnetic bubble is 2µm. To store more information in the given area of the magnetic material, more number of magnetic bubbles is required in the same area. Bubble chips with more than 10 million storage cells are currently used and its read time less than 0.005 sec. The maximum velocity with which the bubble can propagate varies from V_x= 20 to 30 m/s

Bubble Propagation:

Bubble operations used in chip designs are:

- Propagation
- Generation
- Transfer

- Reproduction and
- Destruction

The most basic operation is propagation. The bubble propagation circuit is known as T-bar pattern. The propagation of bubble by T-bar perm alloy pattern is shown in fig 3.19.

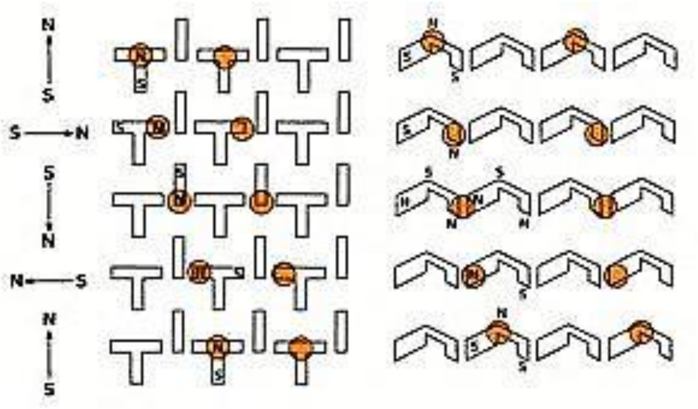


Fig.3.19 Propagation of bubble by T-bar Permalloy pattern

To control the direction of movement, magnetic ‘paths’ are created by deposits of magnetically. Conductive material on the surface of the thin film is in a specific pattern.The in plane drive field producer by this Permalloy T-Bar pattern is used to shift the magnetic bubbles as shown in fig.3.20. Three coils are used to rotate the in plane drive field. The bubble move to the right with the clockwise rotation of the in plane drive field and of the left with the anti-clockwise rotation of the in plane drive field. The drive field rotates in a constant velocity along the material, but the bias field and easy arts are normal to the material. The main drawbacks are

- If required high recording and time for storing and retrieving data.
- If required interface circuits.

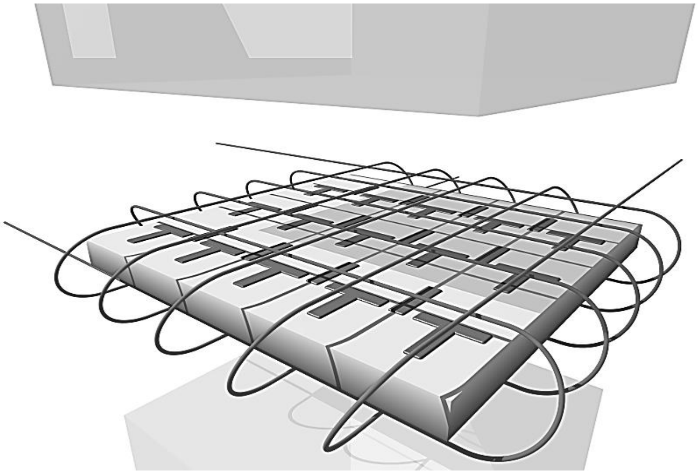


Fig. 3.20 Structure of magnetic bubble memory

The data in the bubble memories are stored similar to other storage devices using the logic 0 (zero) and logic 1 (one) states. There are 3 types of fields acting on the bubble. They are.

1) External magnetic Field or Bias Field (H_B):

A permanent magnet is used to provide the bias field which is necessary to maintain stable bubbles. It is used to reduce the volume (size) of the bubble by exerting a direct force on the bubble wall which is directed towards its centre.

2) Effective domain (bubble) wall field (H_w):

This field is used to reduce the surface area of the bubble wall.

3) Magneto static (stray) field (H_A):

It is used to equalize the overall magnetic surface change.

i.e, Surface charge on the bubble. = Surface charge in the other places of the material.

In the magnetic bubble memory, there is on major loop and 157 minor loops as shown in fig. 3.21 which are arranged from right to left. Each minor loop has 641 bubble sites. Thousands of coded characters may be stored in a single chip.

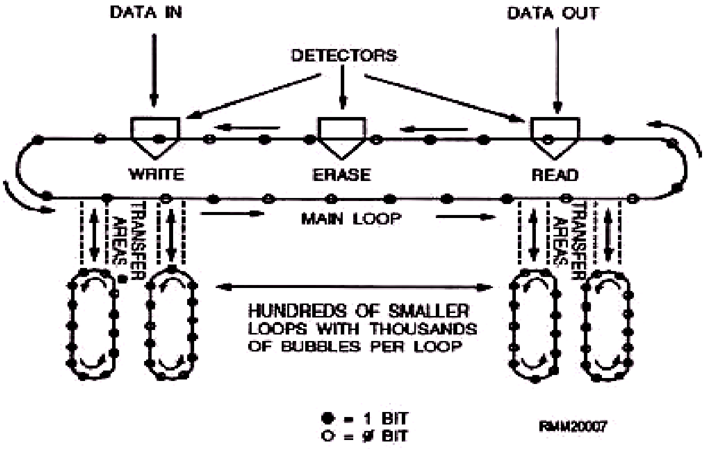


Fig.3.21 Magnetic Bubble Memory

Writing Operation

When a data has to be stored, the bubble from the minor loops are transferred to major loop and it goes to write station. In write station the message is entered and the bubble site again comes to minor loop.

Reading Operation

To read the data from the storage, the bubble from minor loops are transferred to major loop and it goes to read station, then it comes to miner loop. The data can be altered by the erase station, if we need to erase it

Special Features

- It is Non-Volatile
- Bubble sites are moved electronically

- If has high access speed.
- If can store the data even in the absence of electrical power.

It has high storage capacity than magnetic hard disk

3.5. INTRODUCTION TO NANOMATERIALS

Nanotechnology is the term used to cover the design, construction and utilization of functional structures with at least one characteristic dimension measured in nanometers. Nanomaterials are the materials with atoms arranged in nano sized clusters which become the building block of the material. Any material with a size between 1 and 100nm [10^{-9} m to 10^{-7} m] is also called nanomaterials. Such materials and systems can be designed to exhibit novel and significantly improved physical, chemical and biological properties phenomena and processes as a result of the limited size of their constituent particles or molecules, in the range of about 10-9 m to 10-7 m (1 to 100nm). Hence, one classification of nanostructured material and systems essentially depends on the number of dimensions while lie within the nanometer range as shown in Fig.3.22. (a) System confined in three dimensions (i.e, Zero dimension Eg. Quantum dot) (b) System confined in two dimensions (i.e, One dimension, Eg. nanowire, nanotube, nanofiber, etc) (c) System confined in one dimension (i.e, Two dimension, eg. Graphene, Quantum well).

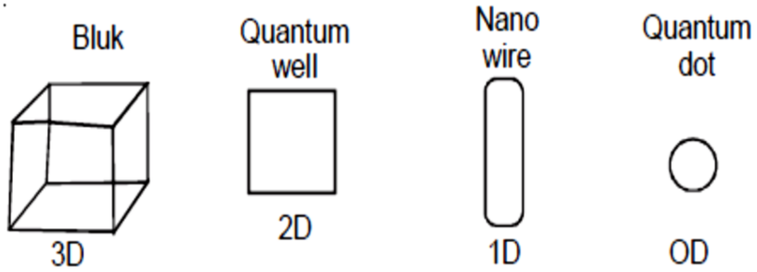


Fig. 3.22 Dimensions of material

Richard Feynman discovered the "bottom up" approach of fabricating materials and devices atthe atomic or molecular scale, possibly using methods of self-organization and self -assembly ofthe individual building blocks. An alternative "top-down" approach is the ultra miniaturizationor etching/milling of smaller structures from larger ones. Both approaches require a means ofvisualizing measuring and manipulating the properties of nanostructures; computer based simulationsof the behaviour of materials at these length scales are also necessary.

The design, fabrication and use of nanostructured systems and the growing, shaping or assemblingof such systems either mechanically, chemically or biologically to form nanoscale architectures,systems and devices. The spin transport in metal, metal-insulator and semiconductorsheterostructures holds promise for the next generation of high-speed, low-power electronic devices.Amongst important spintronic effects already used in practice one can indicate a GiantMagneto Resistance (GMR) in magnetic multilayers &Tunneling Magneto Resistance (TMR) in FMinsulator-FM(FM-I-FM) structure. In addition, spin degree of freedom can be utilized for storinginformation in new forms of memory devices. Aside from the investigation of basic quantumphenomena "artificial atoms" are also promising for applications in high-functionality nanoscaleelectronic and photonic devices such as ultra small memories or high performance laser.

The term "NanoDevice" is indeed a broad term. Here, we will refer to devices at the nanoscale as,using ITRS (International Technology Roadmap for Semiconductors) terminology, emerging researchdevices.

3.6. INTRODUCTION TO NANO DEVICES

Nano Devices are devices that have been developed with nanomaterials mostly of ferromagneticproperty (nano magnets) / semiconducting property (nano semiconductors).

As we know, nanomaterials are materials of ultrafine grain size of the order of 1nm to 100nm; they can be synthesizedwith different dimensions (zero, one, two and three) and with different shapes (powder,rod, wire, tube, pillar, chain, granules, etc).

In order to understand the concepts of nano devices, as a prerequisite, it is necessary that thereaders should have a broad idea about the concepts ferromagnetism, semi conductivity andlaser fundamentals.

3.7.NANO MAGNETS

Nanomagnets are ferromagnets where the magnetic domains are aligned in one dimension, the magnetic dipoles are of nano scale length (10-9m) and are synthesized in different shapes such as dots (Quantum), pillars, wires, tubes, rods, chains, granules etc. Magnetism is an intrinsic physical property associated with the spin of the electrons in a material. Ferromagnetic material exhibit permanent magnetic dipole moment (i.e.,even in the absence of external magnetic field, they have magnetic moment) and it produces a spontaneous magnetization (ie even for a small external magnetic field the amount of magnetization obtained is very large).

Based on process method, nanomagnets are classified as:

Particulate nanomagnets

Geometrical nanomagnets

3.7.1.Particulatenanomagnets

They are granular solids with one or more phases having magnetic properties. There are two methods that make use of the rapid solid technique. In one method, amorphous phase material is first obtained and then it is subjected into necessary heat treatment to get the nano structure. In the second method, the molten melt is first transferred into the block and then it is cooled slowly to obtain the nano structure. In this process, the nanomagnetic material properties can be changed from softest (low coercivity and anisotropy) to the hard magnetic state (high coercivity and anisotropy).

Examples :Nb_3Cu, $CoCrTa$, $Nd_2Fe_{14}B$, $Fe_{73.5}B_9Si_{13.5}$

Layered structure consists of Functional Layer and Seed Layer. Seed layer is first grown with close packed direction with small size grains of CoCr. The functional layer is grown over the textural seed layer with nanostrucural phase of CoCrTa(Fig. 3.23). Characteristics of nanomagnet is high energy (BH)max product, therefore one can store more magnetic energy in a nano magnet. The highest value of energy product required for NdFeB type nano magnet is 450 KJ/m^3. They are used in the manufacturing of devices like cordless power tools, miniaturized ear phones on portable audio devices, high efficiency electro motors.

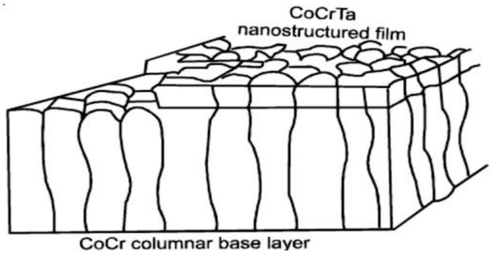


Fig. 3.23 Particulate nanomagnets

3.7.2.Geometrical Nanomagnets

Self assembly

In lithography, a mask is formed on a polymer resist. A pattern of magnetic islands is retained after etching or lift off. On a template of base material, nanoparticles are self assembled through surfactant and ligand structures cover larger surface area. The ligand lengths control the spacing between the nanoparticles. Here again, the energy product is very high.

In the case of ‘shape anisotropy’ where magnetic particles are of normal structure along the field direction and are elongated in a direction to that of the applied field. This shape anisotropy factor is not a desirable one and for this reason AlNiCo which is a geometrical nanomagnet are not considered to be important as that of NdFeB which is a particulate nanomagnet. Self assemblednanopermanent magnets find potential application in MEMs and NEMs.

3.8. MAGNET RESISTANCE

Concepts used in nano devices

1. Ordinary magneto resistance (OMR)
2. Giant magneto resistance (GMR)
3. Tunneling magneto resistance (TMR)

3.8.1.Ordinary magneto resistance (OMR)

It is the property of the material to change the value of its electrical resistance by less than 5% change under the influence of an external magnetic field. It was first discovered by William Thomson in 1856. This can be explained by considering corbino disk as in Fig.3.24. It consists of a two concentric annular rings with perfectly conducting rims. A battery drives the current between the rims, in the absence of external magnetic field. When a magnetic field is applied, a Lorenz force drives a circular component of current, thus increasing the resistance between inner and outer rims and this resistance is called as magnetic resistance.

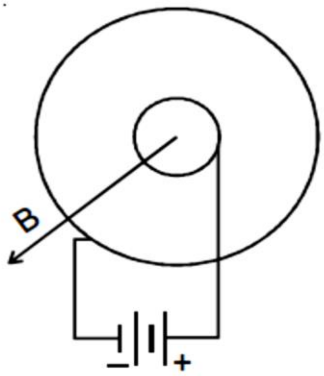


Fig. 3.24Corbinodisc

3.8.2. Giant Magneto Resistance (GMR)

It refers to large change in electrical resistance typically by 50% change due to the changein the magnitude/direction of external magnetic field that occurs in a device that comprisesalternating very thin layers of magnetic metallic elements. Here, the change in electrical resistance is very large compared to other magnetic sensors (where it is only few percent,say 2% to 6%). It is named as giant not because of the size of the device, but for the quantum mechanical magneto resistance effect that leads to large change in electrical resistance in thin layers of magnetic materials. GMR refers to hard disc drive storage technology.

This giant magneto resistive effect was first discovered by the two European researchers,Peter Gruenberg and Albert Fert in the late 1980s and in the year 2007, Nobel prize in Physics was awarded to them for the discovery of GMR.

3.8.2.1 *GMR device (or) GMR heads (or) GMR sensor (or) spin valve*

GMR sensor (or) head (or) device is used to read/write data in a hard disc, by making use of GMR technique.

Structure of GMR sensor (Fig.3.25)

GMR sensor is made up of 4 layers namely,

(i) free layer; (ii) spacer; (iii) pinned layer; (iv) pinning layer ; these layers combine into a single structure.

(i) Free layer :It is the sensing layer of ferromagnetic material and is passed over the surface of the data bits to be read. It is free to rotate in response to the magnetic patterns on the disc. (NiFe) or cobalt.

(ii) Spacer: It is made up of thin layer of non-magnetic metal layer usually made from copper that separates the magnetization of free and pinned layers.

(iii) Pinned layer: A layer of a fixed magnetic orientation by its proximity to the pinning /exchange layer. (Cobalt or / NiFe)

(iv) Pinning (or) Exchange Layer: A layer of strong antiferromagnetic that fixes the pinned layer's magnetic orientation (FeMn).

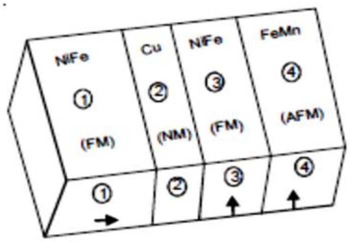


Fig. 3.25 Structure of GMR

(1) Free layer (or) sensing layer (ferromagnetic)

(2) Spacer of non magnetic material

(3) Pinned layer of ferromagnetic material

(4) Exchange layer (or) pinning layer of strong antiferromagnetic material

FM-Ferromagnetic

AFM---Anti Ferromagnetic

NM---Non Magnetic

Working / Mechanism of GMR sensor (or) head (or) device (or) spin valve

The GMR sensor is designed in such a way that it would operate at room temperature and at weak external magnetic field. Here the weak external magnetic field refers to the field due to the bit (1 or 0) on the hard disk that contains up or down spin of the electron.

When such a weak magnetic field (that arises from a bit on a hard disk) passes beneath such a GMR sensor, the magnetic orientation of the free (sensing) layer (1) rotates relative to that of the pinned layer (3) whose magnetic orientation is held by adding the fourth layer (4) of a strong antiferro magnetic material.

The following sequence depicts the activity inside a GMR sensor.

Case 1: Under no external magnetic field (ie) when there is no bit beneath the GMR sensor.

Here the resistance offered is zero as shown in the graph (Fig.3.26a)

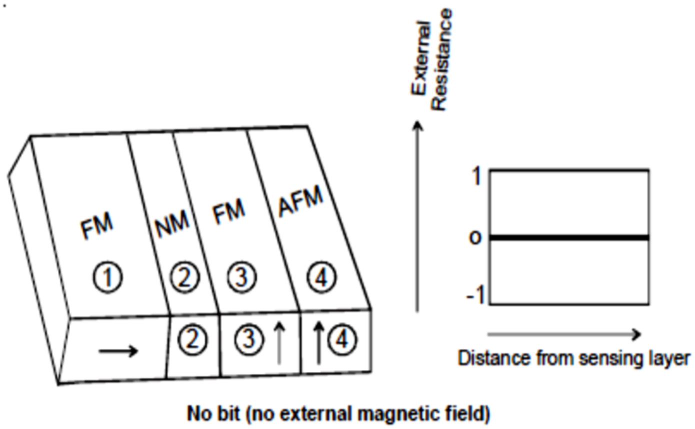
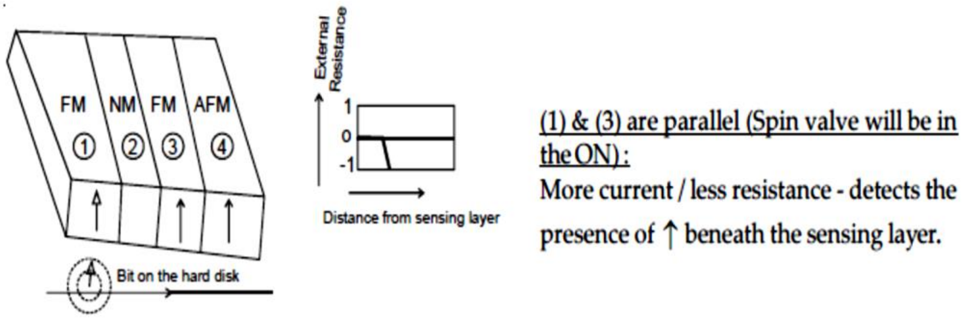


Fig. 3.26 (a) No external magnetic field

Case 2 :As the bit travels exactly under GMR sensor i.e., now the sensing layer is under the weak external magnetic field.



<u>(1) & (3) are parallel (Spin valve will be in the ON) :</u>

More current / less resistance - detects the presence of ↑ beneath the sensing layer.

Fig. 3.26 (b) Weak external magnetic field

Following the domain theory, spin alignment of (1) becomes parallel to that of the bit benath (Fig.3.26c). Hence, the magnetic orientation of sensing layer (1) rotates relative to that of the pinned layer(3). Hence the spin alignment of two adjacent (sensing layer and pinned layer) layers become parallel. Therefore the electron spin from the bit of disk are not scattered by both the sensing layer and pinned layer. Hence the electron spin easily passes through boththe layers. Since, here the scattering is very low, the total resistance is very small (shown in the Fig.3.26b).

Note It can be compared to the case where polarizer and analyzer are parallel to each other, the light can easily pass through.

As the bit moves on, resistance increases since the electron scatters more and hence the current decreases.

Case 3 : As the bit travels away from the GMR sensor

Here the sensing layer will try to orient its spin in a direction opposite to that of the pinned layer. Hence the resistance increases as shown in the graph.

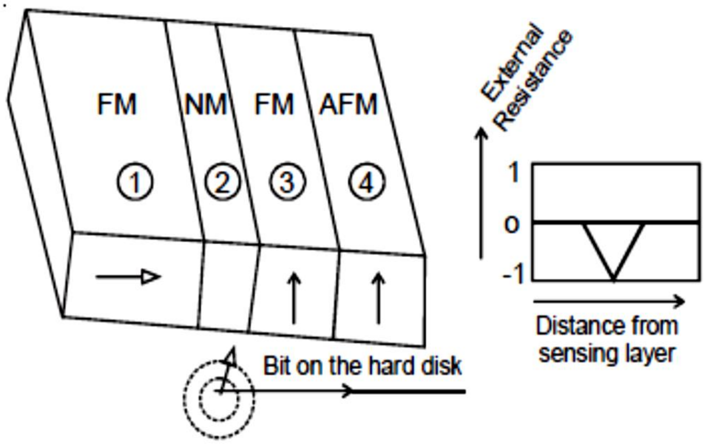
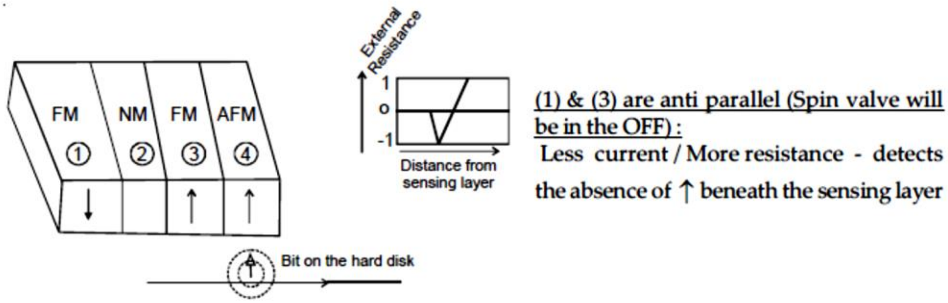


Fig. 3.26 (c) Bit travels away from the GMR sensor

Case 4: As the bit travels far away from the sensing layer of the GMR sensor.

Here, the sensing layer's spin orientation becomes exactly opposite to that of the pinned layer (i.e.) the spin alignment of sensing layer and pinned layer becomes anti parallel (Fig.3.26d).



(1) & (3) are anti parallel (Spin valve will be in the OFF) :
Less current / More resistance - detects the absence of ↑ beneath the sensing layer

Fig. 3.26 (d) Bit travels far away from the sensing layer

Then the spin polarization of the bit will get scattered by the sensing layer and if it is not scattered by the sensing layer it will get scattered by the pinned layer (3) and vice-versa. Since the scattering is more, total resistance is more, shown in the graph.

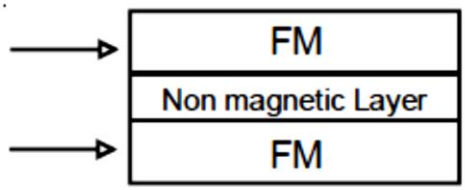
Note: It can be compared to a case when the polarizer &analyser are not aligned parallel to each other and hence the light is blocked. Hence by measuring the total resistance of the GMR device (sensor) head, it is possible to detect the spin orientation of the bit that is beneath the head and on the disk (i.e.) when the total resistance is high, it detects the absence of ↑beneath it. When the total resistance is comparatively less, it detects the presence of ↑.

Since, in the above GMR device, the resistance changes are caused due to the changes in the spin characteristics of the free (or) sensing layer, GMR heads are also known as spin valve, a term coined by IBM. (i.e) GMR sensor acts as a spin switch that can be turned ON which allows the electron spin to pass through hence more current -less resistance; it can be turned OFF –scattering of electrons, thereby the electron spin does not pass through hence less current -more resistance.

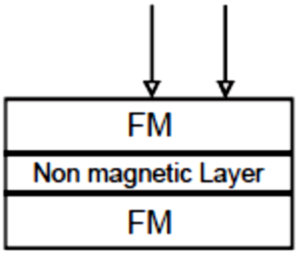
Hence it is named as 'spin valve'. It filters one spin polarization over another spin, hence the name ‘spin valve’. If the layers are not of proper thickness, however coupling mechanism can destroy the GMR effect; hence the layer’s thickness should be optimum and they should be thinner, say several nm than the mean free path of electrons.

External magnetic field can be applied in two ways.

(i) Current is parallel to the plane of layers (CIP) - it offers higher resistance.



(ii) Current is perpendicular to the plane of the layers (CPP) - it offers only less resistance.



3.8.3.Tunneling Magneto Resistance (TMR)

Tunneling Magneto Resistance is the drastic change of tunneling current that passes through a very thin insulating layer which is sandwiched between two ferromagnetic (FM) layers, due to relative changes in the alignment of magnetization of the two ferromagnetic layers. The change in electrical resistance is larger than that in GMR by a factor of 10.

Magnetic Tunneling Junction (MTJ)

It is a device used to read / write data in a non volatile computer memory, namely Magneto Resistive Random Access Memory (MRAM) using TMR technique.

3.8.3.1.*Structure of TMR device / structure of MTJ*

It is very similar to that of GMR device, with the change that here the spacer is a thin insulating layer. It has four layers namely (Fig.3.27a),

1. Sensing layer (Co)
2. Spacer (Al_2O_3/MgO)
3. Pinned layer (Co)
4. Pinning layer (IrMn)

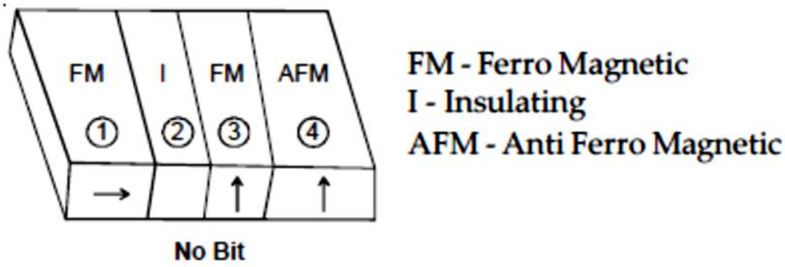


Fig. 3.27 (a) Structure of MTJ

Mechanism

Consider the Magnetic Tunneling Junction, when no voltage is applied to the junction, electrons tunnel in both directions with equal rates.

Under a bias voltage, electrons tunnel preferentially to the positive electrode. [According to classical mechanics, no current can flow through ainsullating barrier but according to quantum mechanics, there is a non vanishing probability for an electron on one side of the barrier to reach the other side].

When the bit of hand disk / CD, passes beneath the sensing layer (1) the spin orientation of the sensing layer rotates relative to that of the pinned layer (3).

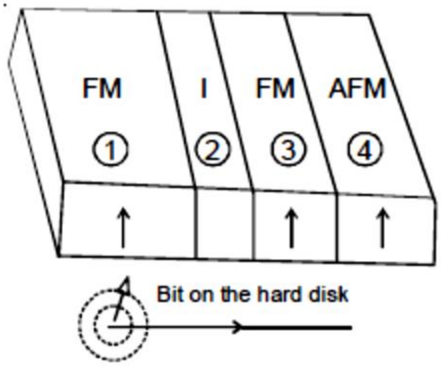


Fig.3.27 (b)

The spin orientation of sensing layer and pinned layers are parallel to each other as in Fig.3.27b. Then, majority of electrons with up spin (↑) of sensing layer (1) will find many empty states of similar orientation in pinned layer (3).

Similarly (↓) electrons of (1) will find many empty states of similar orientation in pinned layer (3). Hence, there is two partial currents tunnel across the insulating layer, one for spin-up ↑ and another for spin-down ↓ that contribute the total current.

Hence large current tunnels through the insulating layer and lowers the overall resistance. Here the spin flip scattering length should be greater than thebarrier width to tunnel across the barrier.

When the bit moves away, the spin orientation of sensing layer (1) becomes anti parallel with that of pinned layer (3) (Fig.3.28d). Here both spin-up ↑ and spin-down ↓ orientations will encounter a bottleneck in either of the two films.

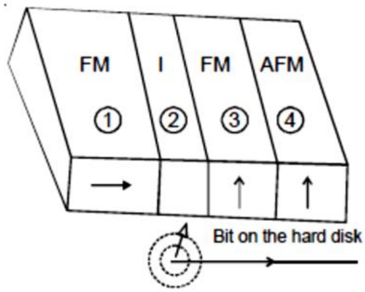


Fig.3.28 (c)

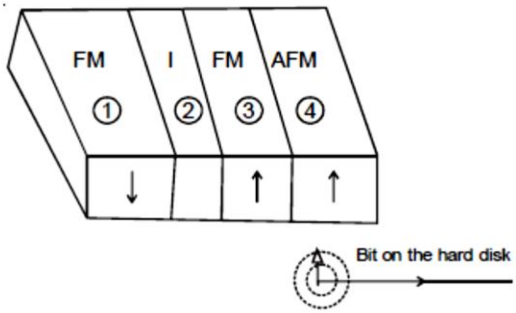


Fig.3.28 (d)

Therefore less empty states will be available in (3) for the spin-down ↓ of (1). Similarly less spin-up ↑ electrons are available in (1) which will occupy the empty states of (3).

Hence the current is small, leading to high total resistance.

Low Resistance: If the two magnetic layers are parallel as shown in Fig. 3.28(c), majority of electrons in 1st FM layer will find many states of similar orientation in the 2nd layer causing a large current to tunnel through and lowering overall resistance. Here, the spin flip scattering length should be greater than the barrier width to tunnel across the barrier.Higher resistance: However, if the two magnetic layers are antiparallel, both spin orientations will encounter a bottleneck in either of the two films, resulting in a high total resistance Fig. 3.28(d). During tunneling, spin orientation is conserved. This technique allows perpendicular recording of data in disk drives, thereby increasing the density of datas to be stored in thegiven area. It exceeds 1 Tera byte / $(\text{inch})^2$. It has led to the emergence of non volatile Magneto Resistive Random Access Memory (MRAM) and reprogrammable magneto logic devices. Non-volatile refers to the computer memory that can retain the stored information even when not powered.

3.8.3.2.*Difference between GMR and TMR*

Even though both GMR and TMR act as a spin valve, the mechanism used for spin filtration is quite different in these two techniques.

GMR	TMR
1. If filters one spin polarization over anotherdue to the changes in the spin characteristicsof the sensing layer, in an externalmagnetic field.	1. If allows both majority and minorityspins due to changes in the tunneling current that occurs due to change inalignment of spins in an external magnetic field.
2. Structures: 4 layers	2. Structures : 4 layers
(i) Sensing layer	(i) Sensing layer
(ii) Spacer - non magnetic conducting layer	(ii) Spacer - insulating layer.
(iii) Pinned - Layer	(iii) Pinned Layer
(iv) Pinning Layer	(iv) Pinning Layer
3. Change in electrical resistance typicallyabout 50%, which is large compared toordinary magneto resistance where thechange is only about 2% to 6%.	3. Change in electrical resistance is largerthan GMR by a factor of 10. 4. During tunneling, spin orientation isconserved.

4. Spin orientation is not conserved. 5. This technique do not lead to recording ofdata in disk drives. 6. It is a hard disk drive storage technology.	5. This techniques allows for recording of data in disk drives, thereby increasing the density of datas stored in the givenarea; it exceeds 1 Tera byte/(inch)^2. 6. It has led to the emergence of Magneto Resistive Random Access Memory (MRAMs) and reprogrammable magnetologic devices. MRAM - is a new type ofnon-volatile memory.

3.9.ProbingNanomagnetic Materials

Magnetic Force Microscopy (MFM) uses a combination of Tapping Mode, Lift Mode (Fig. 3.29) and a magnetized tip to gather information about the magnetic field above a sample. Each sample line is scanned using Tapping Mode to obtain topography data first. This topographic information is stored and retraced with a height offset in Lift Mode, and the magnetic data is then collected. Designed for magnetic force microscopy, these conical-tipped high aspect ratio AFM probes yield a very high force sensitivity and feature a magnetic coating (available in various thicknesses) covering the cantilever and tip.

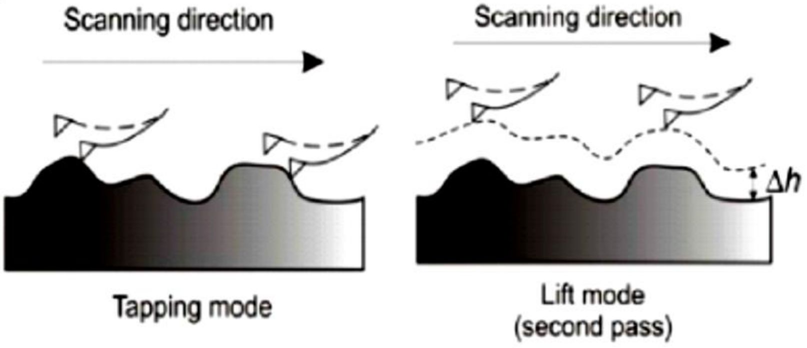


Fig. 3.29 Principle of MFM

PPP(Point Probe Plus)-MFMR AFM probe is our standard probe for magnetic force microscopy providing a reasonable sensitivity, resolution and coercivity. It has proven stable imaging of a variety of recording media and other samples. The force constant of this probe type is specially tailored for magnetic force microscopy yielding high force sensitivity while simultaneously enabling tapping mode and lift mode operation. The hard magnetic coating on the tip is optimized for high magnetic contrast and high lateral resolution of considerably better than 50 nm.

PPP-LC-MFMR AFM probe is coated with a soft magnetic thin film enabling the measurement of magnetic domains in soft magnetic samples. Due to the low coercivity of the tip coating the magnetization of the tip will easily get reoriented by hard magnetic samples.

PPP-LM-MFMR AFM probe is designed for magnetic force microscopy with reduced disturbance of the magnetic sample by the tip and enhanced lateral resolution - compared to the standard PPP-MFMR probe. The force constant of this probe type is specially tailored for magnetic force microscopy yielding high force sensitivity while simultaneously enabling tapping mode and lift mode operation.

SSS-MFMR AFM probe is optimized for high resolution magnetic force imaging. The Super Sharp Silicon tip basis combined with a very thin hard magnetic coating result in an extremely small radius of the coated tip and a high aspect ratio at the last few hundred nanometers of the tip the essential demands for high lateral resolution down to 20 nm in ambient conditions.

When the AFM tip is coated with a magnetic material shown in Fig. 3.30, then the AFM can also measure and create maps of the magnetic structure and characteristics of the sample surface.

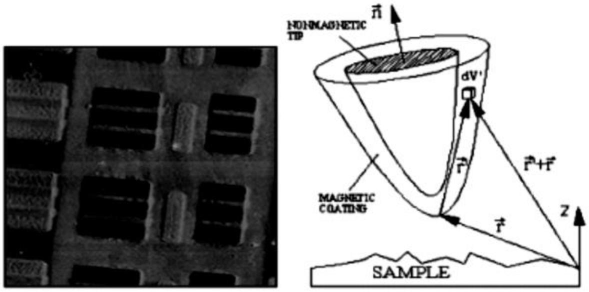


Fig. 3.30 MFM tip having a magnetic coating on a non-magnetic core

3.10. Hard Disk Drive Technology

The main application of hard nanostructures is for the preparation of bonded magnets. Although economically marginal today, MEMS constitute a specific domain of applications, since the magnetic properties of nano structured materials are maintained at small magnet dimensions. Finally it is realized that hard nano structured materials could find applications as high density recording media. Especially used in the fields of magnetic recording, sensors, logic systems, MRAM and spin-electronics.

Bonded Magnets

One can find high performance bonded magnets in our daily lives in various electronic devices, office automation equipment, automotive components and home appliances such as computer hard disk drives, CDs, DVDs. Isotropic powders are most commonly used for the production of bonded magnets. Four RE powder families are commercially used:

(i) isotropic and anisotropic nanostructured Nd-Fe-B, ribbons,

(ii) "lean neo" $Nd_2Fe_{14}B/Fe/Fe_3Bn$ and composites,

(iii) $SmCo_5$ and $Sm_2Co_{17}$and finally

(iv) Sm-Fe-N.

Recording Media

Magnetic data storage constitutes a large potential market for hard magnetic films (Fig. 3.31). While the bit density increases by about 60% each year, the need appears for materials with higher coercivities. Compared to others applications, the amount of material required is much less.

Fig. 3.31 Hard Disk

The media consist of films of typical thickness below 50 nm, themselves made of hard magnetic nanoparticles embedded in a non-magnetic matrix. The use of high anisotropy materials allows the signal to noise ratio to be reduced, while preserving thermal stability. FePt and CoPt films are good candidates for this type of applications, since high corrosion resistance is needed additionally.

MEMS Applications

The application of actuators in micro-electro-mechanical systems (MEMS) requires the miniaturization (100 to 500 microns) of the permanent-magnet components while preserving as much as possible their bulk magnetic properties. In many instances, a severe degradation in coercivity occurs when the magnet size is reduced at such small dimensions. This is particularly true for Nd-Fe-B magnets. The use of nano-size hard magnetic nanostructures allows coercivity to be preserved in micron-sized powders. Screen-printing, bonding and injection molding are well adapted to the fabrication of thick films for MEMS. Most often, isotropic ribbons are used, and the dilution of the powders in the non-magnetic binder limits further the value of the remnant magnetization.

Medical Application

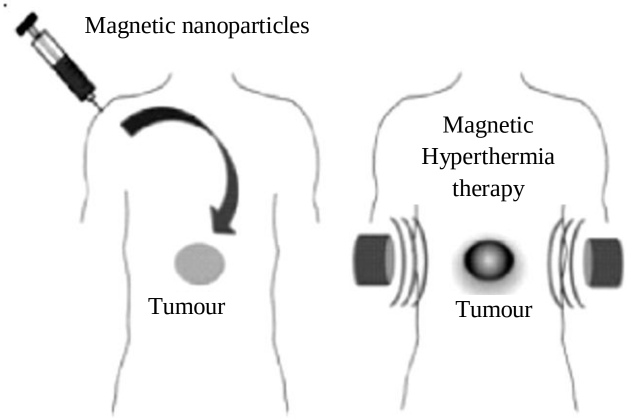


Fig. 3.32Injection of the NMP's suspension into the human body

The main attraction of magnetic nanoparticles from the perspective of drug delivery is the ability to use the magnetic properties to either limit the drug to a particular region using magnetic targeting as shown in Fig. 3.32, or to release the drug remotely. The majority of magnetic nanoparticles used for drug delivery are based on iron oxide or iron-carbon combinations.

Mitoxantrone, doxorubicin, mitomycinC, etoposide, paclitaxel, oxaliplatin and epirubicin have been bound to ironoxide or Fe-C fluids for magnetically targeted cancer treatment.

Magnetic resonance imaging (MRI) aids in diagnosis, research, and treatment of a wide variety of pathologies. Most clinically used nanoparticle contrasts agents are iron-oxide based and is grouped into two categories: Super paramagnetic iron oxides (SPIOs)are between 500 nm and 50 nm, and ultra-small SPIOs (USPIOs), which are smaller than 50 nm.

SPIOs are used for gastrointestinal tract, liver, and spleen imaging and take advantage of the natural clearance of the RES, the transit time through the gut or preferential uptake by specific cells. The most promising uses of SPIOs are to improve the sensitivity of detection and localization of primary and metastatic brain tumors, inflammation and ischemia (insufficient supply of blood to an organ). USPIOs can be used as 'blood pool agents' because their smaller size allows them to remain in the blood stream for longer times. They can be used to assess perfusion (the passage of blood into an organ or tissue) in areas of ischemia and provide information about capillary permeability. They also can be used to study the extent of tumor neovascularity and associated permeability changes.

3.11. Fundamentals of Semiconductors

The electrical properties of a solid material are a consequence of its electron band structure, (i.e.) the arrangement of the outermost electron bands and the way in which they are filled with electrons.

The band that contains the highest energy electrons (or) valence electrons is termed as the valence band; the next higher energy band, which, under most circumstances, virtually unoccupied by electrons, is termed as the conduction band.

The energy corresponding to the highest filled state at zero Kelvin is called the Fermi Energy (E_F). This represents the top-most filled electron energy state of the given material at zero Kelvin.

For semiconductors, all the states in the valence band are completely filled with electrons and there is no overlap between the valence band and the empty conduction band. This gives rise to a narrow energy band gap (<2eV) between the valence band and conduction band. The E_F lies within the band gap (Fig. 3.33).

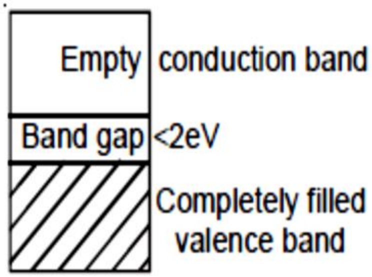


Fig. 3.33 Band Structure for a semiconductor

For electron transition to take place to / from the conduction band with the absorption/ emission of photons respectively, both energy and momentum must be conserved.Although a photon can have considerable energy, its momentum hν/c, is very small; 'h' represents the Planck's constant, ν represents the frequency of radiation and 'c' the velocity of light.Depending on the shape of the band gap and as a function of the momentum K, one can classify the semi conductors as direct band gap semiconductors and indirect band gap semiconductors.

In direct band gap semiconductors(Fig.3.34), the conduction band minimum (Ec) and valence band maximum (Ev) energy occur at the same value of momentum K. Hence, it is possible for direct recombination of electron and hole, emitting a light photon of energy hν. This type of light emission is not possible in indirect band gap semiconductors. Life time or recombination time of charge carriers is very less. Direct band gap semi conductor devices have much higher internal quantum efficiency. They are used to fabricate LED and Laser Diode. Examples for Direct band gap semi conductors are compound semiconductors, formed from III - V group elements and that formed from IV - VI group elements of the periodic table namely, Ga Al As, In Ga As P for (III - V), PbTe, Pb Sn Te for (IV - VI).

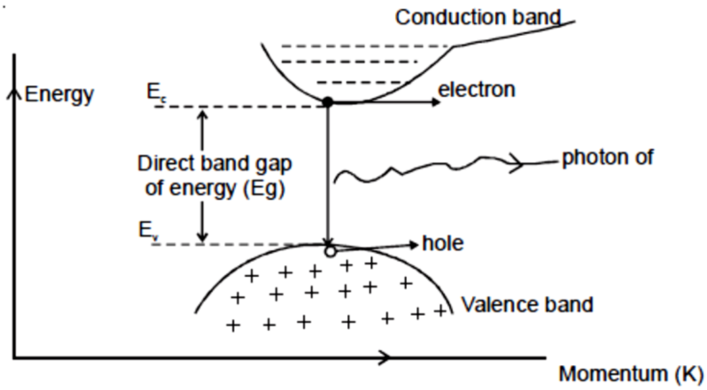


Fig. 3.34 Direct band gap semiconductor

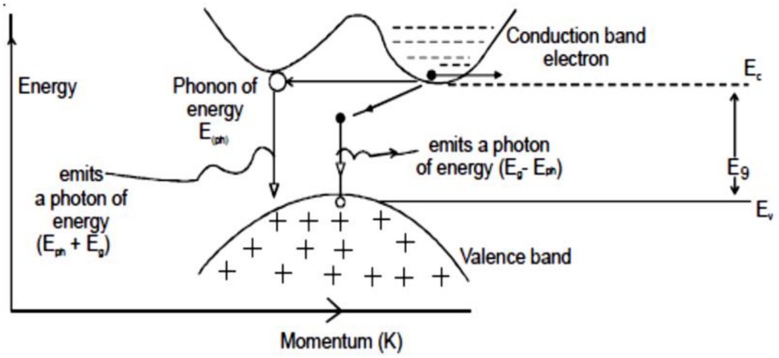


Fig. 3.35 Indirect band gap semiconductor

For indirect band gap semiconductors (Fig.3.35), the conduction band minimum and the valence bandmaximum occur at different values of momentum. To perform band-to-band recombinationand to conserve momentum, it must involve a third particle called phonon (crystal latticevibration), since the photon momentum is very small. Life time of charge carriers is more.Due to longer life time of charge carriers, they are used to fabricate the rectifier diodes andtransistors which are used to make switches and integrated circuits.

Quantum well, quantum wire and quantum dot are nothing but nanostructured semiconductors.The electronic characteristics of these nanostructured semiconductors are closelyrelated to the size and shape of the individual crystal. Generally, smaller the size of thecrystal, larger the band gap and hence greaterthe energy difference between the highestvalence band (EV) and the lowest conduction band (EC). Therefore, more energy is needed toexcite the nanostructured semiconductors and concurrently more energy is released whenthe crystal returns to its resting state. In nanostructured semiconductor lasers, the thicknessof the active medium is around 10 nm. This changes the electronic and optical propertiesdramatically since the dimensionality of the free electron motion is reduced from threedimensions to two dimensions in the case of quantum well and is reduced to one dimensionin the case of quantum wire.

Quantum Dot is a nanostructured semiconductor whose excitons (Electron-Hole) are confined in all three spatial dimensions. Quantum Wire is a nanostructured semiconductor whichconfines the electrons / holes in two spatial dimensions and allow free propagation in the third dimension (i.e.) in a single dimension. Quantum Well is a nanostructured semiconductor which confines electrons / holes in one dimension and allow free propagation of electrons / holes in two dimension. These quantum wells, wires and dots are generally grown by advanced epitaxial techniques such ion implantation, state - of - the - art lithographic techniques, etc.,

3.12 FUNDAMENTALS OF LASER

LASER: Light Amplification by Stimulated Emission of Radiation. Let us see the basic keywordsused in lasers.

3.12.1 Conditions for laser action

Laser action means the amplification of light by stimulated emission of radiation. To get laseraction, it is necessary to have

(i) Stimulated emission.

(ii) There should be population inversion of atoms.

(iii) There should be a stimulating photon.

Population inversion

Stimulated emission of light can be produced through maintaining the population inversion,where there will be more number of atoms unit volume (N_2) in the higher energy level (E_2)than the other lower energy level (say E_1) between which the lasing action takes place; let N1represents the number of atoms / unit volume in E_1.

Then, Population Inversion represents $E_2 > E_1$and $N_2 > N_1$which is possible under non-thermal equilibrium by pumping the atoms from lower energylevel to higher energy level.

Optical Pumping

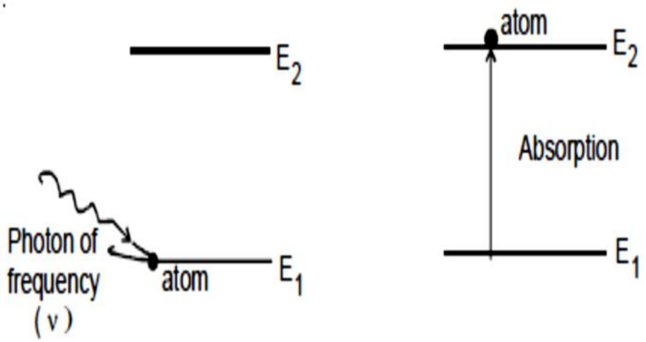


Fig. 3.36 Optical Pumping

The atoms in the lower energy level E_1 are made to go to the upper energy level E_2 bypumping technique, where, by absorbing a photon of frequency $\nu = (E_2 - E_1)/h$, the atom in lowerenergy level E_1, goes to the upper energy level E_2. The rate of absorption depends on the numberof atoms present in the level E_1 and density of photons in the system (Fig.3.36).

Different pumping techniques such as optical pumping, chemical pumping, etc are available.

3.12.2. Spontaneous emission and stimulated emission

Spontaneous Emission

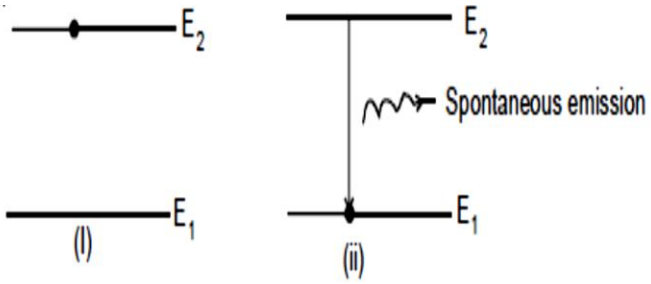


Fig. 3.37 Spontaneous Emission

Stimulated Emission

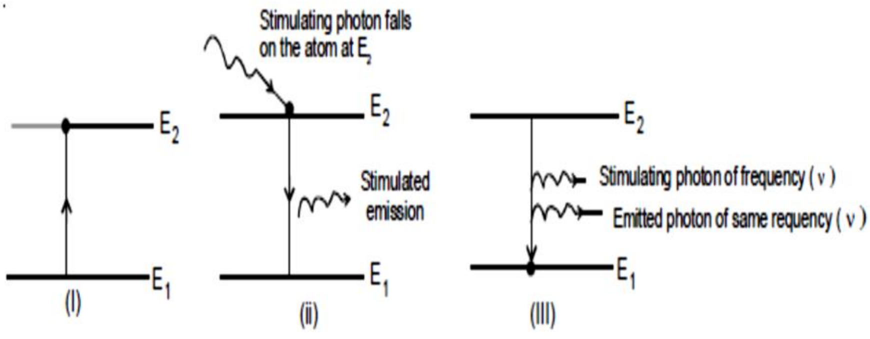


Fig. 3.38 Stimulated Emission

The atom in the exited level can come to the ground level either by spontaneous emission or bystimulated emission.

In spontaneous emission (Fig. 3.37), the atom in energy level E_2 make the transition to the lower energy levelE_1 spontaneously without the help of any external agent and it emits a photon whose energy isequal to (E_2-E_1). The rate of emission depends on the number of atoms present in the higher energylevel.

Characteristics of photons emitted by spontaneous emission are polychromatic radiation,low intensity, incoherent, large divergence.

In stimulated emission (Fig.3.38), the atom in higher energy level E_2 can make a transition to thelowerenergy level E_1 with the help of an external photon of frequency, $(E_2- E_1)/h$,where the stimulatingphoton and the emitted photon are in same phase with the same energy and they travel in thesame direction. Characteristics of photons emitted by stimulated emission are monochromaticradiation, high intensity, coherent and high directionality

Active Medium

Active medium is a medium where the laser action takes place. The medium may be a gasmedium / a liquid medium / a solid medium. The solid medium refers to an insulating crystal/ a semi conductor. (eg) Solid Laser - Ruby laser (Al_2O_3 crystal doped with Cr+ ions); Semiconductor laser, Gas laser - He-Ne laser, CO_2 laser.

Optical amplification is produced in this active medium and the medium acts as an opticalresonator to provide the necessary optical feedback.

3.13 EINSTEIN'S COEFFICIENTS, RELATION BETWEEN SPONTANEOUS AND STIMULATED EMISSION PROBABILITY

3.13.1 *Excitation*

The Bohr's atom model expresses that the electrons are distributed in the permitted energy levels/orbits. The lowest energy level that occupied by electrons is ground state/lower energy state of the atom.

If the electron belongs to lower energy state absorbs the energy to reach the next higher energy level/state is known as Excitation /absorption. Then the energy required for excitation of atom is $E_2 - E_1 = hv$ where E_2 is the energy of excited state of electron in a atom and E1 the energy of ground state of electron in the same atom. This process is called excitation. The rate of excitation is directly proportional to number of atoms (N_1) in energy state E_1 and the energy density (Q) of incident radiation.

$$N_{ab} \infty N_1Q$$

Therefore, the number of excitations occurring per unit time is given by $N_{ab} = B_{12} N_1Q$

where $B_{12}$is a proportionality constant.

3.13.2. *Spontaneous emission*

Spontaneous emission is a continuous process. Since, the life time of excited electron at higher energy state is less (in the order of nano seconds), the electrons from higher energy could release the energy that gained during excitation by some mechanism and make the whole atom to get the excitation. The energy released will be $E_2 - E_1 = hv$

If the energy released is continuous in the form of radiation, which does not follow unique wave length then it must be due to spontaneous emission, The rate of spontaneous emission is directly proportional to the number of atoms in the excited energy state (N_2).

$$N_{sp} \infty N_2$$

Hence the number of transitions per second is given by $N_{sp} = A_{21}N_2$ (2)

Where

A_{21} is proportionality constant. This process is a **downward transition**.

3.13.3. Stimulated emission

Stimulated emission is a controlled de excitation. If the energy released in de excitation is controlled in the presence of some energy which is used for the excitation. This situation causes controlled emission of radiation following the unique wave length known as stimulated emission.

The rate of transition is directly proportional to the number of atoms in the upper energy level and the energy density of incident radiation.

$$N_{st} \infty N_2Q$$

The number of transitions per second$N_{st} = B_{21}N_2Q$ (3)

Where$B_{21}$is proportionality constant. This process is a **downward transition**.

The proportionality constants A_{12}, B_{12} and B_{21} are known as Einstein's coefficients A and B.

Under equilibrium condition, the number of downward and upward transitions per second is equal.

ie., $N_{sp} + N_{st} = N_{ab}$ (4)

Substituting from the eqns 1, 2 and 3, in eqn 4 we have,

$$A_{21} N_2 + B_{21} N_2 Q = B_{12} N_1 Q \qquad (5)$$

Rearranging eqn 5 we have

$$B_{12} N_1 Q - B_{21} N_2 Q = A_{21} N_2$$

$$Q(B_{12} N_1 - B_{21} N_2) = A_{21} N_2$$

$$Q = A_{21} N_2/(B_{12} N_1 - B_{21} N_2) \qquad (6)$$

Dividing numerator and denominator by $B_{21} N_2$, we have

$$Q = A_{21}/B_{21} \text{ x } 1/ (B_{12}/B_{21})(N_1/N_2) - 1 \qquad (7)$$

On substituting $N_1/N_2 = e^{h\nu/kT}$ in eqn (7) we have

$$Q = A_{21}/B_{21} \text{ x } 1/ (B_{12}/B_{21})(e^{h\nu/kT}) - 1 \qquad (8)$$

Planck's radiation formula for energy distribution is given by

$$Q = 8h\pi\nu^3/c^3(e^{h\nu/kT}) - 1 \qquad (9)$$

Comparing the eqns (8) and (9),we have

$$B_{12}/ B_{21} = 1$$

$$B_{12} = B_{21} \qquad (10)$$

and $A_{21}/ B_{21} = 8h\pi\nu^3/c^3$ (11)

Conclusion

The spontaneous emission is more predominant than the stimulated emission. The LASER light is due to stimulated emission. Therefore stimulated emission should be greater than spontaneous emission. To achieve this, population inversion is required.

The equation (11) gives the relation between spontaneous emission and stimulated emission coefficients. Since the ratio is proportional to ν^3, the probability of spontaneous emission increases with the energy difference between the two states.

3.14.INJECTION LASER DIODE (ILD)

Injection Laser Diode is basically laser source that makes use of double hetero junction directband gap semi conductors as the active medium in order to achieve laser light of high quantumefficiency and high power output.

Hetero junction means a direct band gap p-n junction formed by two different semiconductingmaterials with different band gap energies. The material on one side of the junction differs fromthat on the other side of the junction.

Double hetero junction has two such hetero junctions.

Efficient wave guide structure, small beam divergence, high coherence, monochromaticity,carrier and optical confinement, high power output with low threshold current is possible withsuch hetero junction semi conductors.

3.14.1 Structure of ILD

It uses double hetero junction semi conductors. There are three layers, one is the central activelayer and the other two are the two adjacent layers. Active layer is of p-type material, usually$Ga_{(1-y)}Al_yAs$ of thickness 0.1 to 0.3 μm and of low energy band gap. The adjacent layers are madeup of n-type $Ga_{(1-x)}Al_xAs$ of high energy band gap. Here x is not equal to y and they determine theband gap of alloys. The adjacent layers are also known as injection layers/confinement layers, (Fig.3.39).

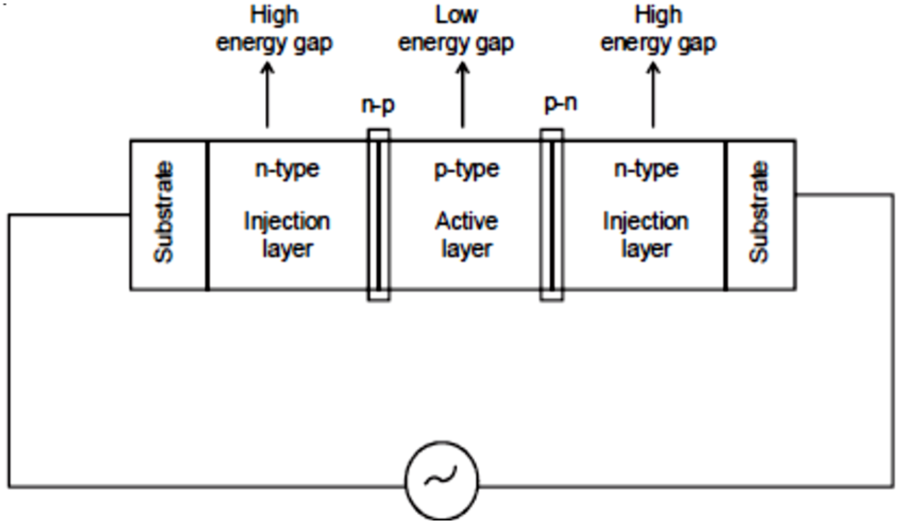


Fig.3.39 Structure of Double Hetro Junction Laser

n-type : $Ga_{(1-x)}Al_xAs$

p-type : $Ga_{(1-y)}Al_yAs$

3.14.2.Working of ILD

When drive current (I) is given to ILD, the excess electrons are moved from Ec of adjacent layer toE_c of active layer in order to achieve population inversion, a necessary condition to be satisfied forlaser action to start. Electron transition occurs in the active layer from E_c to E_v, since only lessenergy has to be crossed in the active layer. Due this transition, electron-hole recombination takes place in the active layer; thereby photon is released from the active layer. This photon can thenstimulate the release of coherent photons. Optical confinement that is necessary for amplificationis obtained by choosing $n_1>n_2$, where n1 is the refractive index of the active layer and n_2 is thatof the adjacent layer.The photon released in the active layer moves and gets totally internally reflected on hitting the n-p and p-n junction, since n1 > n2. Number of times it gets amplifiedand finally it results in the emission of light in 800 - 900 nm wavelength range, (Fig.3.40).

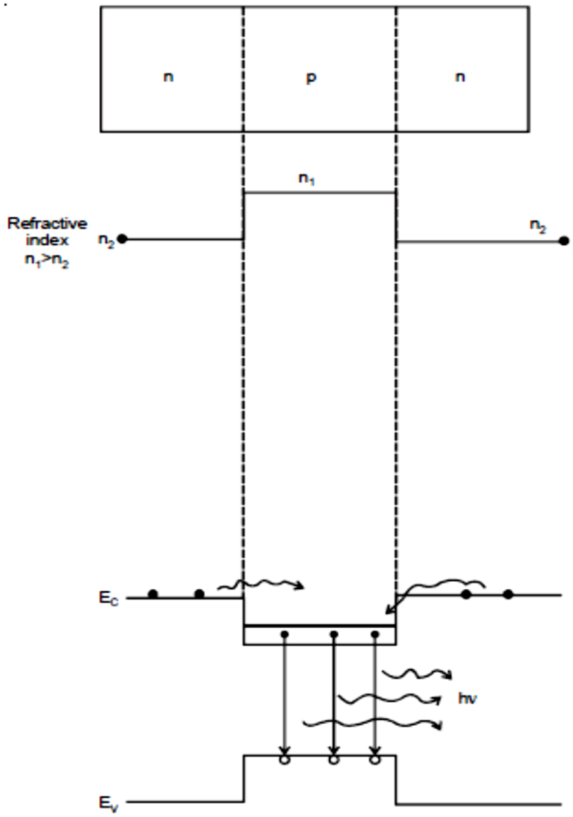


Fig. 3.40 Working of ILD

3.14.3.Merits of ILD

1. High directivity, coherent radiation, monochromatic, high output power.
2. Acts as narrow spectral width source.
3. Coupling efficiency is high hence used to couple laser light into the optical fiber.
4. Used for long distance optical communication at higher bit rate > 200 Mb / sec.
5. Reduces chromatic dispersion.

3.14.4. Demerits of ILD

1. Shorter life time; 10 times expensive than LEDS.
2. Wavelength of the output laser light is governed by the semiconductor band gap.
3. More temperature dependent.
4. In order to maintain population inversion, high power consumption is needed and it consequently increases the temperature of the laser.
5. Due to increase in temperature, (i) Efficiency of the laser decreases rapidly as the electronpopulations are smeared out through the wide band structure of available states. (ii) As thethreshold (J_{th}) current density depends on temperature, J_{th} also increases consequently.

Note: Threshold current density represents the minimum current density of the active mediumthat is needed to start and sustain laser action in such a way that the optical gain inside the laseris greater than or at least equal to the sum of the loss and useful laser output.

$$J_{th}(T) = J_{th}^{o} e^{\frac{T}{T_o}}$$

J^0_{th} - Threshold current density at 0°C.

T - Temperature of the active layer

T_o - Parameter which determines the relative temperature sensitivity of the device.

T_o - Parameter which determines the relative temperature sensitivity of the device.

This threshold current density also changes with the age of the laser. Output wavelength isgoverned by the semi conductor band gap. The problems in ILDs can be overcome to some extentby replacing the bulk semiconductor with a nanostructured semi conductors namely quantumwell, quantum wire and quantum dots.

3.15.QUANTUM WELL LASERS(QWL)

In a normal double hetero structure laser diodes (ILD), the active layer thickness is about0.1 - 0.3 μm. Hence the active region behaves like a 3 Dl box, where the electrons (or holes) arespread over a relatively wide energy range with a small electron density at the band gap edge.The density of states depends on the value of energy.

But in a single quantum well laser (SQW), the active medium is a quantum well of thickness, say 50 - 100Å. Single Quantum Well laser has a single active region. Due to the smaller thickness ofthe active region, the active region behaves like a two dimensional thin film. In a SQW, theelectron's movement is confined to two dimension and this lead to the development of quantizedenergy states within the SQW, called sub-bands. Energy difference between these sub-bands is afunction of thickness of quantum well, hence precise wavelength of light to be emitted can easilybe selected by properly choosing the well width.

Since, the electron is confined in one dimension, the electrons are spread over a very small energyrange, but there is abrupt increase of density of states (DOS) at the band gap edge and is independentof energy of electron. Due to the two dimensional DOS, the electrons and holes are crowdednear the edges of conduction / valence band respectively resulting in population inversion stateand hence increase of probability of stimulated emission, (Fig. 3.41).

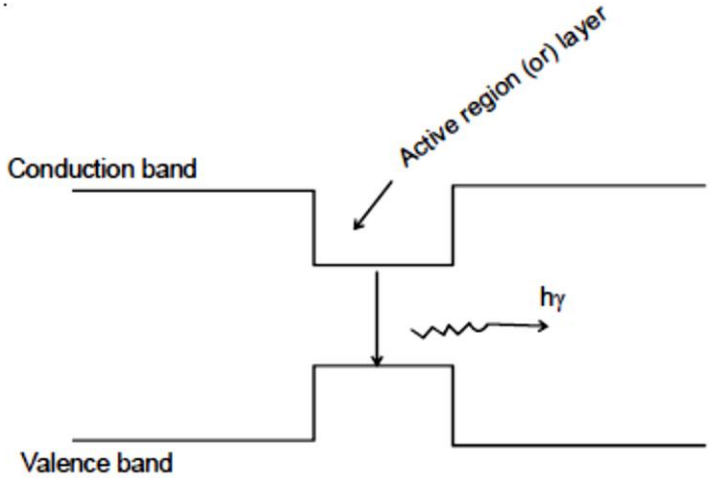


Fig. 3.41 Band structure of Single Quantum Well (SQW) Laser.

3.15.1. Drawback of SQW Laser

Due to extreme narrowness of the gain region, optical confinement is very poor and hence the lossis very high. To rectify this, Multiple Quantum Well Lasers (MQWL) has been fabricated.

3.15.2. Multiple Quantum Well Lasers

MQWL correspond to multiple active regions. Each active layer is an individual laser. Activelayers are separated by barrier layers. We know that a SQWL consists of sub-bands and in aMQWL, a number of SQWLs are brought in close proximity with each other,(Fig.3.42). Hence sub-bands ofthese SQWs interact and as more SQWs are brought up together, these sub-bands start to frommini-bands. They are different from the intrinsic bands of bulk semi conductor.

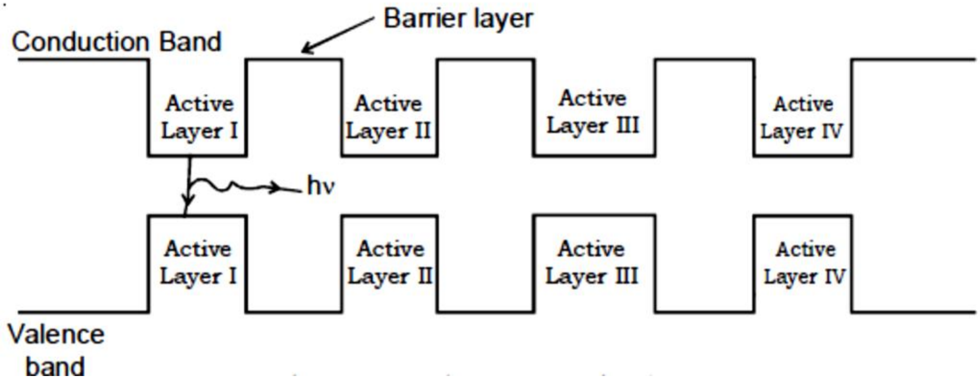


Fig.3.42 Band structure of MQWL

The stack structure of QWs is referred to as 'Super-Lattice', distinguishing it from the crystallattice of the underlying semiconductor layers.

3.15.3. Merits of MQWLs

i. Better optical confinement with very low Jth.

ii. Less temperature dependence.

iii. Narrow line width.

iv. Higher modulation speed.

v. Lower frequency chirp. (Frequency chirp means time dependant phase shift of optical pulses that will lead to optical broadening).

3.15.4.Drawback of Quantum Laser (or) Quantum Well Laser

In order to get the laser output in the far IR range, one requires very low photon energy. But in QWlaser, this photon energy is still close to the band gap of bulk semiconductor; hence it is difficult toobtain laser light in the far IR, using QWs.

3.16. QUANTUM CASCADE LASER (QCL)

It is basically an Injection Laser Diode (ILD), where the bulk semiconductor is replaced by thenanostructured Multiple Quantum Wells(MQW) that act as the active medium and emit laserlight in the Infra red (IR) and in far IR region of the electromagnetic spectrum.

3.16.1. Working of QCL

The structure of QCL is the similar to that of ILD, with the change that n-p-n semiconductors arenanostructured (refer the diagrams of ILD). Initially QWs of slightly different band width andhence different energy states are chosen similar to that as in ILD. In QCL, when drive current isapplied, the electrons cascade down a series of steps emitting a sequence of phonons - quantum oflattice vibrations, hence the name Quantum Cascade Laser Fig. 3.43 (a) & (b).

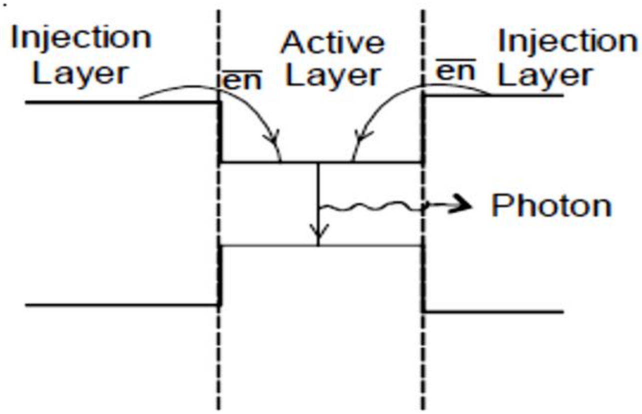


Fig. 3.43 (a) Injection laser Diode

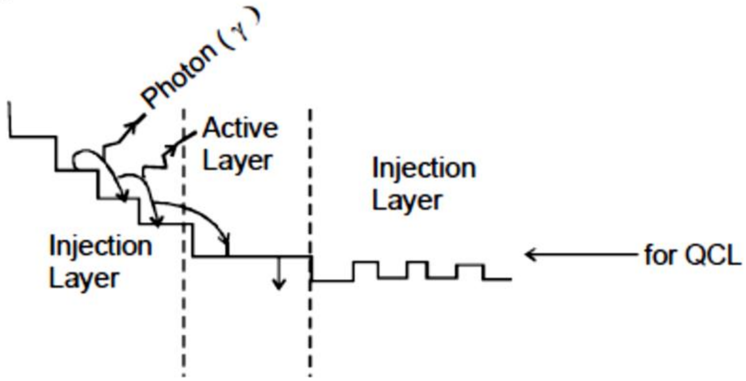


Fig. 3.43 (b) Quantum Cascade Laser

Energy levels of QCL, before the application of bias voltage is shown below:

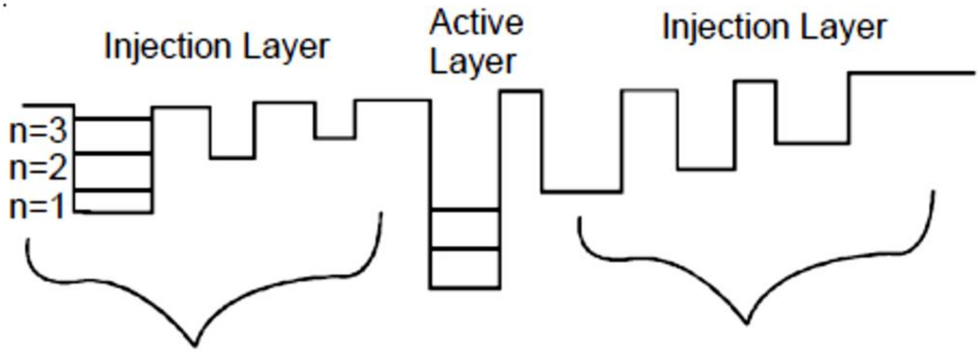


Fig. 3.44 Energy levels of QCL, before the application of bias voltage

Energy levels of QCL, after the application of bias voltage is shown below:

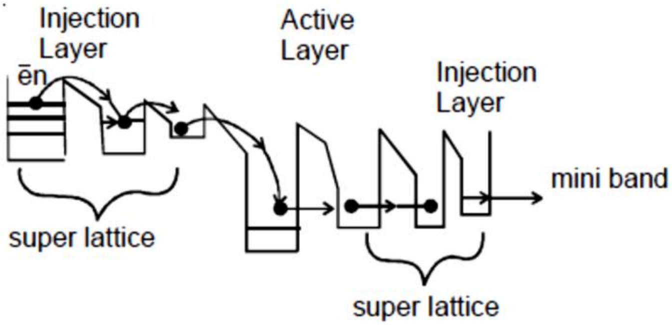


Fig. 3.45 Energy levels of QCL, after the application of bias voltage

Here, the energies of QWs are tuned by adjusting the thickness of each QW so that they matchwith the bias voltage. Hence, the applied bias voltage brings the energy states of these QWs intoresonance creating super-lattice with the required mini-bands.

The fig. 3.45 explains the case when electrons fall from n = 3 to n = 2 and they fall to n=1 state.They excite the miniband of 1st QW and move to the next QW through the active region andphotons are released.Recently developed QCL, has laser output at 3.1μm wavelength at 20k and3.6μm wavelength at room temperature.

3.17. Differences between injection laser diode and quantum lasers are listed below:

ILD	QWL / QCL
1. Active medium - bulk semiconductors. 2. Size of the active medium -0.1 to 0.3μm 3. Active region behaves as a 3D box 4. Electrons / holes spread over wideenergy range with small DOS at the bandgap edge. 5. DOS depend on the energy of the	1. Active medium - nano structured semiconductors. 2. 50Å - 100Å 3. Active region behaves as a 2D thin film. 4. Electron's energy is spread over asmall range with increased DOS at theband gap edge.

electron. 6. Thershold current density is 800A/cm^2 7. Output wavelength is governed by semiconductor band gap;hence control of output wavelength isnot possible. 8. High power consumption; more temperature dependant. 9. Downward transition from conduction band to valence band leads to emission of laser light – interband device. 10.They are interbandbiploar device.	5. DOS is independent on energy ofelectron. 6. Here is is 60 A/cm^2 -which is easy toobtain at room temperature itself. 7. Output wavelength depends on thicknessof nano structured semiconductorhence precise required wavelength oflight can be made to be emitted. 8. Reduced operating current and are less temperature sensitive. 9. Downward transition between conduction band states of different confined QWs - intraband device. 10.They are intraband unipolar device

3.10.3 Merits of QCL

i. High optical power output.
ii. Tuning is possible to get the required wavelength λat far IR.
iii. Possibility of room temperature operation.
iv. Lower operating current.
v. Faster operating speed.

3.10.4 Demerits of QCL

When QCL is operated continuously at room temperature, it generates heat. As a result, electron's energy is wasted and there is a possibility of exciting even from the upper lasing state itself before making a downward transition.

3.10.5 Applications

i. Spectroscopic applications such as gas monitoring, pollution monitoring in atmosphere.
ii. Space communication system, medical diagnostics such as breath analysis, etc.

Just to know

Maiman

On May 16, 1960, at Hughes' Malibu, California, laboratories, Maiman's solid-state pink ruby laser emitted mankind's first coherent light, with rays all the same wavelength and fully in phase. Maiman documented his invention in Nature and published other scholarly articles describing the science and technology underlying his laser.

Maiman had begun conceptualizing a solid-state laser design even before he undertook the maser project at Hughes. Moving the microwave frequency of masers up the electromagnetic spectrum 50,000-fold to the frequency of light would require finding a feasible lasing medium and excitation source and designing the system.

Richard Feynman

Richard Feynman gave a 1959 talk which many years later inspired the conceptual foundations of nanotechnology. The American physicist Richard Feynman lectured, "There's Plenty of Room at the Bottom," at an American Physical Societymeeting at Caltech on December 29, 1959, which is often held to have provided inspiration for the field of nanotechnology. Feynman had described a process by which the ability to manipulate individual atoms and molecules might be developed, using one set of precise tools to build and operate another proportionally smaller set, so on down to the needed scale. In the course of this, he noted, scaling issues would arise from the changing magnitude of various physical phenomena: gravity would become less important, surface tension and Van der Waals attraction would become more important.

QUESTIONS

1. Discuss Ferromagnetism and explain Ferromagnetic Domain theory – Weiss theory.
2. Briefly explain different types of energy involved in domain growth with suitable sketches.
3. Describe the formation of Hysteresis loop using domain wall movement. and hence classify the magnetic materials as hard and soft on the basis of Hysteresis loop.
4. Describe the Formation, Propagation, Read/Write operation and special features of the magnetic Bubble Memory.
5. What are the classifications of nanomagnetic materials? Explain the different nanomagneticmaterials with suitable illustrations.
6. What is GMR? Explain the structure and mechanism of GMR head with a neat diagram.
7. What is TMR? Explain the structure and mechanism of TMR head and its necessary TMR diagram.
8. Explain relation between spontaneous and stimulated emission probability. Derive an expression for Einstein's Coefficient
9. Describe the structure and working principle of ILD with its necessary diagram. Write itsmerits and demerits.
10. Describe
    (i) Structure and working of Quantum cascade laser (QCL) with diagram
    (ii) Merits and demerits of QCL.
11. Give the applications of nanomagnets in hard disk drive technology.

SOLVED PROBLEMS

3.1 A paramagnetic material has a magnetic field intensity of 10^4 A/m. If the susceptibility of the material at room temperature is 3.7×10^{-3}, calculate the magnetization and flux density in the material

Solution

$$\chi = \frac{M}{H} \text{ and } B = \mu_0 (M + H)$$

Hence $M = \chi H$

$$= 3.7 \times 10^{-3} \times 10^4$$
$$= 3.7 \times 10$$
$$= 37\ Am^{-1}$$

Hence magnetization = 37 Am^{-1} (Answer)

Flux density $B = \mu_0 (37 + 10^4)$

$$= (4\Pi \times 10^{-7} \times 10037)$$

$$= (4\Pi \times 10^{-7} \times 10037)$$

$$= 126179.43 \times 10^{-7}$$

$$= 0.0126\ wb/m^2$$

3.2 The saturation magnetic induction of nickel is 0.65 wb/m^2. If the density of nickel is 8906 kg/m^3 and its atomic weight is 58.7, calculate the magnetic moment of the nickel atom in Bohr magneton.

Solution

The number of nickel atoms/m^3 is

$$N = \frac{density \times Avagadro\ number}{atomic\ weight}$$

$$= 8906 \times 6.025 \times 10^{26}$$

$$= 9.14 \times 10^{28} / m^3$$

When χ is very large

$$B_S = \mathrm{N}\mu_0\mu_m$$

$$B_S = 0.65\ wb/m^2$$

$$N = 9.14 \times 10^{28} / m^3$$

$$\mu_0 = 4\Pi \times 10^{-7}$$

$$\mu_0 = 4\Pi \times 10^{-7}$$

$$\mu_m = \frac{B_S}{N\mu_0} = \frac{0.65}{9.14 \times 10^{28} \times 4\Pi \times 10^{-7}}$$

$$= 5.66 \times 10^{-24} Am^2$$

$= 0.61\mu_B$ Bohr magneton

$= 0.61\mu_B$

ie. magnetic moment of the nickel = 0.61 μ_B (Answer)

3.3 A magnetizing field of 1600A/M produces a magnetic flux of 2.4×10^{-5} wb in an iron bar of cross-sectional area 0.2 cm^2. Calculate permeability and susceptibility of a bar.

Solution

We know, magnetic flux

$$\phi_B = BA$$

$$B = \phi_B / B$$

Here $\phi_B = 2.4 \times 10^{-5}$ wb and A = 0.2 $cm^2 = 0.2 \times 10^{-4}\ m^2$

$$B = \frac{2.4\times 10^{-5}}{0.2\times 10^{-4}} = 1.2\, wb/m^2$$

$$= 1.2\, NA^{-1}M^{-1}$$

Magnetic permeability is expressed by the equation

$$\mu = \frac{B}{H}$$

$$\mu = \frac{1.2}{1600} = 7.5 \times 10^{-4}\, NA^{-2}$$

$$\chi_M + 1 = \mu_r = \frac{\mu}{\mu_0}$$

$$\chi_M = \frac{7.5\times 10^{-4}}{4\times 3.14\times 10^{-7}} - 1$$

= 596.13 (Answer)

3.4 The magnetic susceptibility of a medium is 948 × 10^{-4}. Calculate the permeability and relative permeability.

Solution

$\mu_r = 1+\chi_M$, $\mu_r = 1+948\times 10^{-11} \approx 1$ or > 1, If μ is the permeability of the medium and μ_0 that of vacuum,

$$\mu_r = \frac{\mu}{\mu_0} \quad or \quad \mu = \mu_r\mu_0 \text{ , } \mu_0 = 4\Pi\times 10^{-7} \text{ , } \mu = 1\times 4\Pi\times 10^{-7}$$

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF PHYSICS

UNIT – IV - PHYSICS FOR ENGINEERS-SPHA1101

UNIT 4

THERMAL PHYSICS

4.1 INTRODUCTION TO THERMAL PHYSICS

The study of thermal physics is important because of its application in various fields of engineering.**Thermodynamics** is the branch of physics that deals with heat and temperature, and their relation to energy, work, radiation, and properties of matter. The behavior of these quantities is governed by the four laws of thermodynamics which convey a quantitative description using measurable macroscopic physical quantities, but may be explained in terms of microscopicconstituents by statistical mechanics. Thermodynamics applies to a wide variety of topics in science and engineering, especially physical chemistry, chemical engineering and mechanical engineering, but also in fields as complex as meteorology.

Thermodynamic process are used in mechanical, electrical,civil and in chemical engineering applications.

Mechanical Engineering

In designs of internal and external combustion engines, refrigeration and air-conditioning plants, heat-exchangers, coolers, condensers, furnaces, pre heaters.

Electrical Engineering

In designing cooling system for motors, transformers and generators

Civil Engineering

Thermal Physics

Control of heat transfer in dams, structures, buildings and tunnels

Chemical Engineering

The application of heat transfer is in freezing, boiling, evaporation and condensation processes

A thermodynamic system is a group of material and/or radiative contents. Its properties may be described by thermodynamic state variables such as temperature, entropy, internal energy, and pressure. The simplest state of a thermodynamic system is a state of thermodynamic equilibrium, as opposed to a non-equilibrium state. A system is defined as quantity of matter or a region in space chosen for study. Everything external to the system is surrounding. Thermodynamic system and surrounding is always separated by the boundary. In a **closed system**, no mass may be transferred in or out of the system boundaries. The system always contains the same amount of matter, but heat and work can be exchanged across the boundary of the system. Whether a system can exchange heat, work, or both is dependent on the property of its boundary. But, An **isolated system** is more restrictive than a closed system as it does not interact with its surroundings in any way. Mass and energy remains constant within the system, and no energy or mass transfer takes place across the boundary.

4.2 LAWS OF THERMODYNAMICS

- An **open system** which is where energy and matter can be exchanged between a system and its surroundings.
- A **closed system** where only energy can be exchanged between a system and its surroundings, not matter.
- An **isolated system** where no energy or matter is exchanged between a system and its surroundings. A truly isolated system is rare.

Thermal Physics

The First Law of Thermodynamics: The First Law of Thermodynamics, also known as the Law of the Conservation of Energy, says that energy cannot be created or destroyed, it can only change form. *Energy comes in a variety of different forms.*

Energy is not created or destroyed; it simply changes from one form to another. Turning on a light switch does not create energy, it simply converts electrical energy into radiant energy (light) and thermal energy (heat).

$$\Delta \mathbf{U = Q - W + (Energy)}$$

Here, ΔU is the total change in internal energy within a system, Q is the heat exchanged between the system and its surroundings, and W is the work done to or by the system.Within the First Law are three related concepts – work, heat, and internal energy. **Heat** is the transfer of thermal energy between two systems. **Work** is the force that transfers energy between a system and its surroundings. By producing work either within a system or outside of it, you create heat. Then there's **internal energy**, which is all the energy within a system. When heat, work, and internal energy interact together, energy is transformed.

The Second Law of Thermodynamics:

The Second Law of Thermodynamics, also known as the Law of Increased Entropy, says that over time the state of disorganization or **entropy** in a system will always increase

The Third Law of Thermodynamics

The Third Law of Thermodynamics says that a perfect crystalline structure at absolute zero temperatures will have zero disorder or entropy. However, if there is even the smallest hint of imperfection in this crystalline structure, then there will also be a minimal amount of entropy. Regardless, this law allows us to understand that as the entropy of a system approaches a temperature of absolute zero, the entropy present within a system decreases.

The Zeroth Law of Thermodynamics: The Zeroth Law of Thermodynamics says that if two systems are in thermal equilibrium with a third system, then the first two systems are also in thermal equilibrium with one another.

- If System A is in balance with System C
- And System B is in balance with System C
- Then System A and System B are also in balance with each other.

4.3. THERMAL EXPANSION OF SOLIDS

It is a common observation that nearly all substance expand on heating and contract on cooling.The change in temperature of a body causes expansion or contraction of that body. Most of the substances expand on heating and contract on cooling. This is called thermal expansion.

Types of expansion

A solid substance can undergo three types of expansion, expansion in length, expansion in area and expansion in volume.

(i).Expansion in length is known as linear expansion

(ii).Expansion in area is known as superficial expansion

(iii).Expansion in volume is known as cubical expansion

Coefficient of linear, superficial and cubical expansion of a solid:

4.4.1. Coefficient of linear expansion of a solid

The coefficient of linear expansion of a solid is the increase in length of unit length of the solid when its temperature is raised by 1K. It is denoted by the letter α

Thus if a rod having length L_1 at T_1K is heated to a temperature T_2 K its length increase to L_2

$$\alpha = \frac{L_2 - L_1}{L_1(T_2 - T_1)} = \frac{Increase\ in\ length}{Original\ length\ \times Rise\ in\ temperature}$$

4.4.2. Coefficient of superficial expansion of a solid

The coefficient of superficial expansion of a solid is the increase in area produced per unit area of the solid when its temperature is raised by 1K. It is denoted by the letter β

Thus if a solid having area A_1 at T_1 K is heated to a temperature T_2 K its area increase to A_2.

$$\beta = \frac{A_2 - A_1}{A_1(T_2 - T_1)} = \frac{Increase\ in\ area}{Original\ area\ \times Rise\ in\ temperature}$$

4.4.3. Coefficient of cubical expansion of a solid:

The coefficient of cubical expansionof a solid is the increase in Volume per unit volume of the solid for 1K rise of temperature. It is denoted by the letterγ.

Thus if a solid having volume V_1 at T_1 K is heated to a temperature T_2 K, let V_2 volume at T_2 K. then

$$\gamma = \frac{V_2 - V_1}{V_1(T_2 - T_1)} = \frac{Increase\ in\ volume}{Original\ volume\ \times Rise\ in\ temperature}$$

4.5. THERMAL EXPANSION OF LIQUIDS

When a liquid is heated in a container, heat flows through the container to the liquid, which means that the container expands first, due to which the level of the liquid falls. When the liquid gets heated, it expands more and beyond its original level. We cannot observe the intermediate state. We can only observe the initial and the final levels. This observed expansion of the liquid is known as the apparent expansion of the liquid. If we consider the expansion of the container also and measure the total expansion in volume of the liquid, then the expansion is termed as the absolute expansion of the liquid

Absolute and apparent expansion of a liquid.

A liquid has always been taken in a container and so when heat is given to the vessel; both the liquid and the container expand in volume.

The observed expansion of the liquid is only relative to the container. If we ignore the expansion of the container what we obtain is only apparent expansion.

To find the absolute expansion of the liquid we should take into account the expansion of container.

4.5.1. Coefficient of apparent expansion of a liquid

It is the observed increase in volume of unit volume of the liquid per degree Kelvin rise of temperature. It is denoted by γ_a.

If V_1 and V_2 be the observed volumes of the liquid at temperature T_1 K and T_2 K respectively.

$$\gamma_a = \frac{V_2 - V_1}{V_1(T_2 - T_1)} = \frac{Increase\ in\ volume}{Original\ volume\ \times Rise\ in\ temperature}$$

4.5.2. Coefficient of real or absolute expansion of a liquid

It is the real increase in volume of unit volume of a liquid per degree Kelvin rise of temperature. It is denoted by γ_r

If V_1 and V_2 be the observed volumes of the liquid at temperature T_1 K and T_2K, then

$$\gamma_r = \frac{V_2 - V_1}{V_1(T_2 - T_1)} = \frac{Real\ Increase\ in\ volume}{Original\ volume\ \times Rise\ in\ temperature}$$

The units of coefficients of apparent and real expressions of the liquid are per ^{o}C or per K.

4.6. EXPANSION JOINTS

An expansion joint or movement joint is an assembly designed to safely absorb the heat induced expansion or contraction of a pipeline, duct or vessel. It is used to absorb vibration, to hold parts together.

They are commonly found between sections of buildings, bridges, sidewalks, railway tracks, and settlement of pipeline, ships and other structures.

Building forces, concrete slabs and pipelines expand and contract due to warming and cooling from seasonal variation or due to other heat sources. Before expansion joint gaps were built into these structures and they would crack under the stress induced.

We know the young's Modulus

$$Y = \frac{Longitudinal\ stress}{Longitudinal\ strain} \quad --\rightarrow (1)$$

$$Longitudinal\ stress = \frac{Force}{Area} = \frac{F}{A} \quad --\rightarrow (2)$$

We know that Coefficient of thermal expansion of a solid

$$\alpha = \frac{dl}{l\theta} \quad --\rightarrow (3)$$

From this, we can write longitudinal strain $= \frac{dl}{l} = \alpha\, \theta$

$$\text{Young's Modulus } Y = \frac{F/A}{\alpha\, \theta}$$

$$\frac{F}{A} = Y\alpha\, \theta$$

$$F = Y\alpha\,\theta\ \ A \qquad --\rightarrow (4)$$

From equation (4) we can see that if the area is less than the force required to restore the materials to its original position is less. Suppose if the area is large then the restoring force should also be more, which is quit impossible. Hence, to avoid this problem, while constructing a large area of beams, gaps are provided & these gaps are called expansion joints.

Examples

- Expansion joints are provided while lying the railway lines.
- Expansion joints are provided even in the constructions of buildings. However the joints area well packed and are not visible.

4.7 BIMETALLIC STRIPS

Definition

A Bimetallic strip means a strip made up of two thin metals pined together. These two metals have with different coefficient of thermal expansion. It is operated in the principle of different metal have different coefficients of expansion.A bimetallic strip is used to convert a temperature change into mechanical displacement. The strip consists of two strips of different metals which expand at different rates as they are heated, usually steel and copper, or in some cases steel and brass. The strips are joined together throughout their length by riveting, brazing or welding.

Fabrication

Let us consider two metals

i. Brass of higher coefficient of thermal expansion

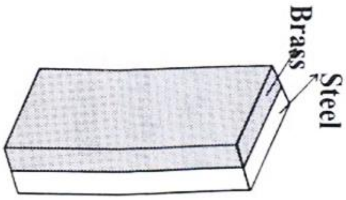
ii. Steel of lower coefficient of thermal expansion

While heating

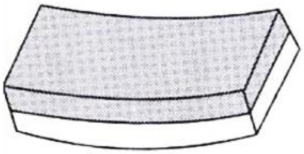
When bimetallic strip is heated then the strip will start expanding and therefore the brass, which has large co-efficient of thermal expansion expands more than the steel and hence the bimetallic strip bends like arc as shown in figure.

While Cooling

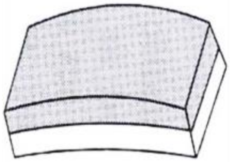
When bimetallic strip is cooled then the strip will start contracting and therefore the brass, which has large co-efficient of thermal expansion contract more than the steel and hence the bimetallic strip bends like arc as shown in figure.



Bi metallic Strip



Heating

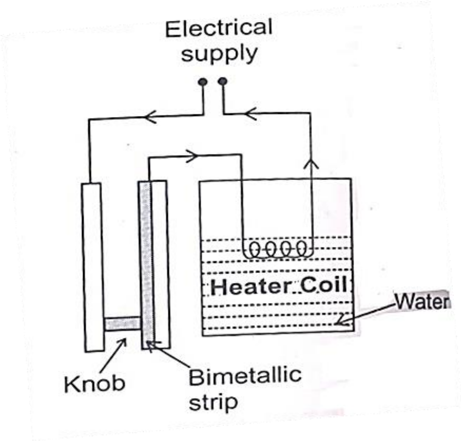


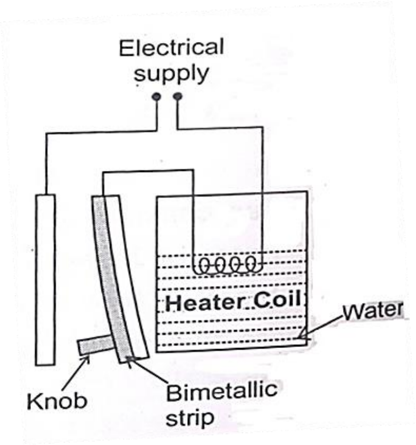
Cooling

Applications

Bimetallic strips are commonly used to water heaters as temperature controller as detailed below.

A water heater connected to the bimetallic strip as shown in figure. At room temperature the bimetallic strip remains straight; the circuit is in closed condition.

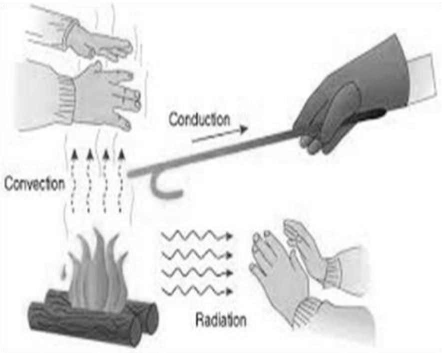


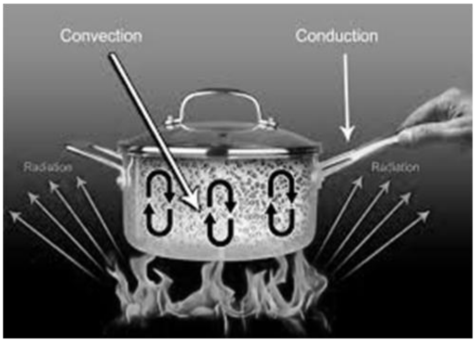


When the power supply is switched on due to increase in temperature of the water, the heater coil becomes hot and in turn the bimetallic strip also becomes hot. Now due to thermal expansion, the bimetallic strip starts bending and at a particular temperature the knobs get detached and hence, the circuit becomes open. Hence the bimetallic strip acts as a temperature controller of water heater.

4.8. TRANSFER OF HEAT ENERGY

Heat is one of the forms of energy. It is transferred from one place to another place by three different modes such as (a) Conduction (b) Convection (c) Radiation.





4.8.1. Conduction

Thermal conduction is defined as the transport of energy due to random molecular motion across a temperature gradient. It is the process in which the heat is transferred from hot end to cold end without the actual movement of the particles.

Thermal conductivity refers to the ability of a given material to conduct/transfer heat. It is generally denoted by the symbol ‘k’ but can also be denoted by ‘λ’ and ‘κ’. The reciprocal of this quantity is known as thermal resistivity. Materials with high thermal conductivity are used in heat sinks whereas materials with low values of λ are used as thermal insulators.**Thermal conductivity** of materials is temperature dependent. The reciprocal of **thermal conductivity** is called thermal resistivity. Metals with high **thermal conductivity**, e.g. copper, exhibit high electrical conductivity.

Examples

When one end of the metal rod is heated the molecules at the hot end vibrate with higher amplitude and transmit the heat energy from one particle to other adjacent particle and so on. At the same time each individual particle remains in their mean position of equilibrium.

4.8.2. Convection

It is the process in which the heat is transmitted from hot end to cold end by the actual movement of the particles.Thermal convection involves heat transfer in a convective manner. Convection involves transfer of any form of energy by the modes of transfer from one place to another, (cases which involve movement) this mainly happens in boilers, where the water nearby the furnace gets heated while the water that remains at the top won’t receive that much heat

Example

When some potassium permanganate is added to the beaker of water and heated. We can find that the lower region become warm first and become less dense.

Hence it moves up and the more dense cold water comes down. Now this cold water is heated and becomes less dense so that it moves towards upward direction.

4.8.3. Radiation

It is the process in which heat is transmitted from one place to another directly without the necessity of the intervening medium.

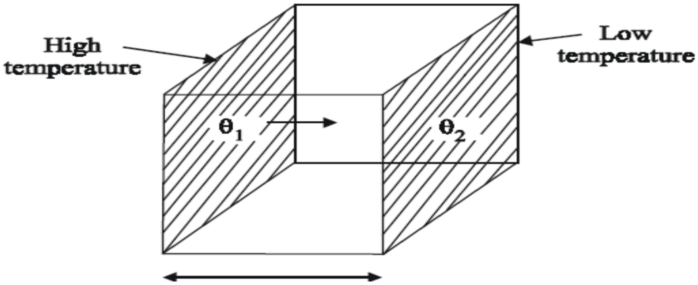
Radiation emitted by such a body due to its temperature is called thermal radiation. All bodies not only emit such radiation but also absorb it from their surroundings. If the rate of absorption and emission from a body become equal, the body is said to be in thermal equilibrium with its surroundings.

Examples

Though there is no material medium between the earth and sun, the radiation from the sun reaches us. Here the energy is sent out in the form of radiant heat waves. The properties of heat radiation are similar to light radiation.

4.9. HEAT CONDUCTIONS IN SOLIDS- THERMAL CONDUCTIVITY

Different material conducts heat differently and so it is essential to know about the conducting power of materials so called thermal conducting power of a material so called thermal conductivity.



It can be defined as the quantity of heat conducted per second through unit area of the materials when unit temperature gradient is maintained.

Let two forces of a materials A and B be separated by a distance x . Let the face A be kept at a higher temperature θ_1K and face B be kept at lower temperature θ_2K. Let the fall of the temperature be uniform and same throughout the material from A and B. If Q is the quantity of heat conducted across the faces then it is directly proportional to the area (A), temperature difference, time of flow (t) and inversely proportional to the distance (x).

$$Q \quad \alpha \quad \frac{A\,(\theta_1 - \theta_1)t}{x}$$

$$Q \quad = \quad \frac{K\,A\,(\theta_1 - \theta_1)t}{x}$$

Where k is the constant called the thermal conductivity of the materials

$$K \quad = \quad \frac{Q\,x}{A\,(\theta_1 - \theta_1)t}$$

If A= 1 m^2;$(\theta_1 - \theta_1) = 1K$ and t= 1s

$$K \quad = \quad Q$$

4.9.1.Temperature Gradient

The quantity $\frac{(\theta_1 - \theta_1)}{x}$ represents the rate of fall of temperature with respect to the distance or thickness which is known at temperature gradient.

4.9.2. Thermal diffusivity (h)

It is defined as the ratio of thermal conductivity to the thermal capacity per unit volume.

$$h = \frac{Thermal\ conductivity}{Thermal\ capacity} = \frac{K}{\rho S}$$

4.9.3. Specific hear capacity

It is defined as the amount of heat required to raise the temperature of unit mass of the substance through one Kelvin.

$$S = \frac{Q}{m\theta} \qquad J\,Kg^{-1}\;K^{-1}$$

4.9.4. Rate of Rise / loss of heat

The rate of heat raise or less with respect to the time is given by

$$\frac{dQ}{dt} = MS$$

4.9.5. Coefficient of Thermal Conductivity

Thermal conductivity is the ability of a material to conduct heat. It is similar to electrical conductivity with the only difference being the entity is heat and not electricity. Thermal conductivity **(λ or k)** is the capacity of the body to conduct or spread heat.

Definition: The coefficient of thermal conductivity is defined as the amount of heat conducted per second, normally across unit area of cross section, maintained at unit temperature gradient.

$$Q = -\;KA\frac{d\theta}{dx}\,t$$

where,

Q = Heat supplied, t = Time, T= Temperature, K = Thermal Conductivity, A = Area of cross-section of the material, x = Thickness of the material, dQ/dt = Rate of heat conduction

dϴ/dx = Temperature gradient

4.10 THERMAL CONDUCTIVITY FOR GOOD CONDUCTORS- FORBES METHOD- THEORY AND EXPERIMENT

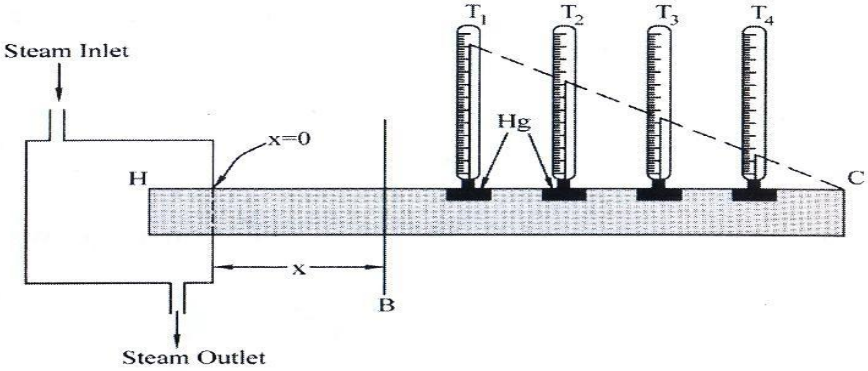
This is the one of the earliest method to find the absolute thermal conductivity of metals.

AIM: To measure the thermal conductivity of good conductors such as Aluminium, Copper, Brass, etc. by following Forbes method.

ExperimentalSetup:

The experimental set up is shown in Fig. 1. One end of the experimental rod in connected to a Heating Chamber. There is a shield to avoid radiation and direct heating of other portion of the rod from the heater. A copper rod has one end held at 100◦C by means of a steam jacket. At steady state, a temperature gradient is established along the rod that can be measured using thermocouple. Analysis of the results enables the thermal conductivity k of the rod to be determined.

Theory and experiments



Consider a long road. This rod is heated at one end and a steady state is reached after some time.

The amount of heat flowing per seconds across the cross-sections B at the distance x from the hot end $Q = KA \left(\frac{d\theta}{dx}\right)_B --\rightarrow (1)$

This amount of heat flowing across the section B is equal to the heat lost between B and C, therefore

The amount of heat lost per seconds by radiation by the rod beyond the section 'B'

$$= \int_B^C m \times s \times \left(\frac{d\theta}{dt}\right)$$

Since Mass = volume ×Density then Mass of the element $= (A\,dx)\,\rho$

$$\text{Heat lost per seconds } = \int_B^C (A\,dx) \times \rho \times s \times \left(\frac{d\theta}{dt}\right) --\rightarrow (2)$$

$Where\ \left(\frac{d\theta}{dt}\right)$- Rate of fall of temperature of the element and S specific heat capacity of the rod

$$\begin{pmatrix} Amount\ of\ heat\ flowing\ per\ seconds \\ across\ the\ section\ at\ the\ point\ B \end{pmatrix} = \begin{pmatrix} Heat\ lost\ by\ radiation\ by\ the \\ rod\ beyond\ the\ section\ B \end{pmatrix}$$

$$KA \left(\frac{d\theta}{dx}\right)_B = \int_B^C (A\,dx)\,\rho\, s \left(\frac{d\theta}{dt}\right)$$

$$K = \frac{\rho s \int_B^C \frac{d\theta}{dt}\,dx}{\left(\frac{d\theta}{dx}\right)_B} --\rightarrow (3)$$

The Experiments consists of two parts

Thermal Physics

1. Static experiment to find $\left(\frac{d\theta}{dx}\right)_B$
2. Dynamic experiments to find $\left(\frac{d\theta}{dt}\right)$ and $\int_B^C \frac{d\theta}{dt}\, dx$

1. Static experiment

The specimen metal is taken in the form of a long road. One end of this rod is headed up to six (or) seven hours by a steam chamber. The rod has a series of holes into which thermometers are fitted. These thermometer record temperatures (T_1,T_2,T_3,T_4)at different points along the rod.

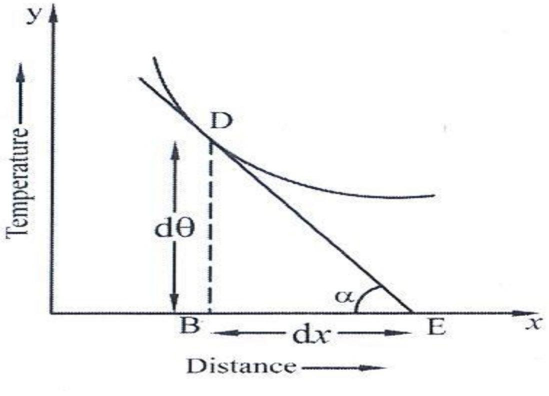
When the steady state is reached, the temperature showed by the thermometers of the road and their respective distance (x) from the hot end are noted

A graph is drawn between the temperature (θ) and distance $(x)$from the hot end.

The value of $\left(\frac{d\theta}{dx}\right)_B$ obtained by drawing tangent to the curve at a pointB.

If the tangent makes an angle α with the x axis, then from the greaph

$$\left(\frac{d\theta}{dx}\right)_B = \frac{BD}{BE} = \tan\alpha \quad --\rightarrow (4)$$



2. Dynamic experiments

A piece of the original rod is heated to the same temperature as that of the hot end in the static experiments. The heated piece of the rod suspended in air.

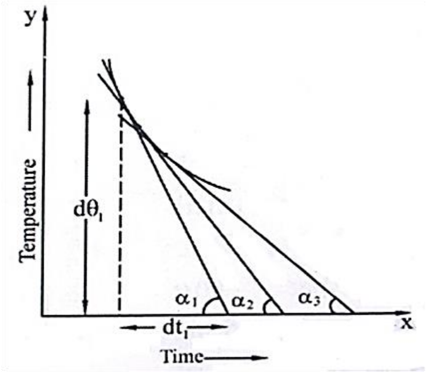
Now it is allowed to cool. Its temperature is noted at regular intervals of time by a thermometer placed in a hole at the centre.

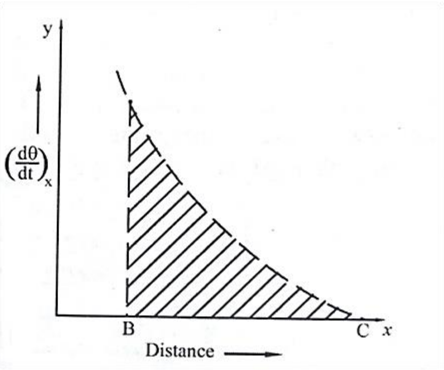
A graph is drawn between time along x axis and temperature along y axis as shown in figure. Different tangents are drawn for corresponding temperature indicated by the thermometer T_1,T_2,T_3, and T_4 in the static experiments.

Now we get the $\left(\frac{d\theta}{dt}\right)_{x_1} = \tan\alpha_1$, $\left(\frac{d\theta}{dt}\right)_{x_1} = \tan\alpha_2$ ….etc.

Rate of cooling at that corresponding distances beyond the section in static experiment is obtained.

With the above data a third curve is plotted taking x from the hot end beyond the section B along x- axis and $\left(\frac{d\theta}{dt}\right)_{x_1}$, $\left(\frac{d\theta}{dt}\right)_{x_1}$ etc. along y axis as shown in figure.





The area of the shaded portion of the curve will corresponding to $\int_B^C \frac{d\theta}{dt}\,dx$, where B is the point chosen in the static experiments.

$$Area\ of\ the\ shaded\ portion = \int_B^C \frac{d\theta}{dt}\, dx \quad --\rightarrow (5)$$

Substituting equation (4) and (5) in equation (3)

The thermal conductivity $K = \dfrac{\rho \times S \times Area\ of\ the\ shaded\ portion}{\left(\frac{BD}{BE}\right)}$

The experiment is repeated by choosing the B at different distance from the hot end and average value of K is determined.

4.11. THERMAL CONDUCTIVITY OF A BAD CONDUCTOR – LEE’S DISC METHOD

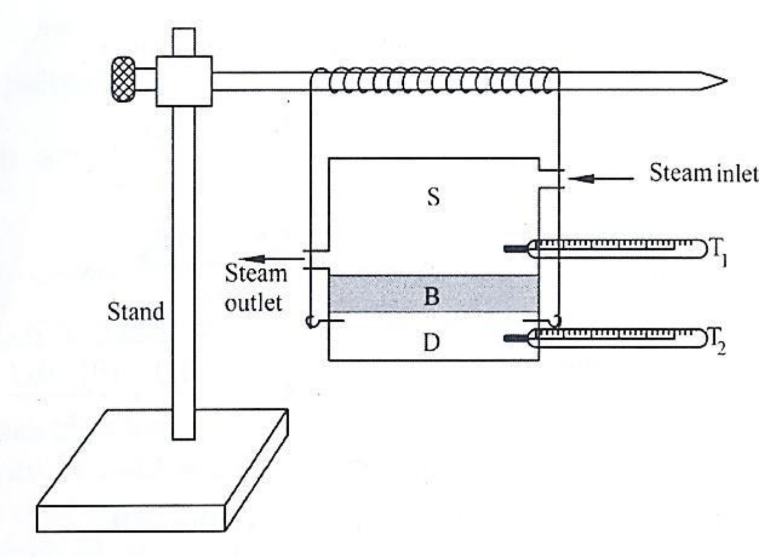
Principle

In the steady state, the quantity of heat flowing across the bad conductor in one second is equal to the quantity of heat radiated in one second from the lower face area and edge area of the metal disc in the Lee’s disc method.

Description: The given bad conductor (B) is shaped with the diameter as that of the circular slab (or) disc ‘D’. The bad conductor is placed in between the steam chamber (S) and the slab (or) disc (D), provided the bad conductor, steam chamber and the slab should be of same diameter. Holes are provided in the steam chamber (S) and the slab (or) disc (D) in which the thermometers are inserted to measure the temperatures. The total arrangement is hanged over the stand as shown in figure.

Working

Steam is passed through the steam chamber till the steady state is reached i.e., the thermometer show constant temperature. Let the temperature of the steam chamber (hot end) and the disc or slab (cold end) beθ_1 and θ_2 respectively.



Calculation:

Let ‘t’ be the thickness of the bad conductor (B), ‘m’ is the mass of the slab, and ‘S’ be the specific heat capacity of the slab. ‘r’ is the radius of the slab and ‘h’ be the height of the slab, then,

Amount of heat conducted by the bad conductor *per second* = $\frac{K\,A\,(\theta_1 - \theta_2)}{t}$

Since the Area of cross section is = πr^2

Amount of heat conducted per second = $\frac{K\,\pi r^2\,(\theta_1 - \theta_2)}{t}$ ——→ (1)

The amount of heat lost by the slab *per second*

= m ×S× Rate of cooling

The amount of heat lost by the slab *per second*

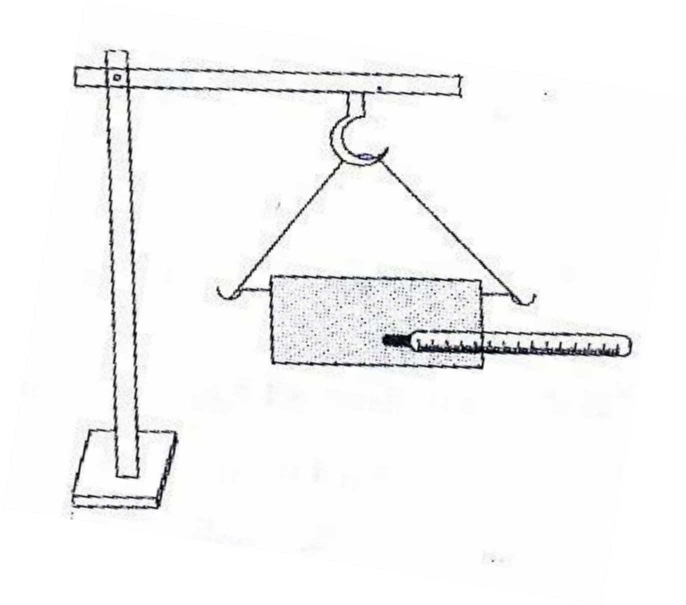
= m × S× R_c ——→ (2)

Under steady state

∴We can write eqn. (1) = eqn. (2)

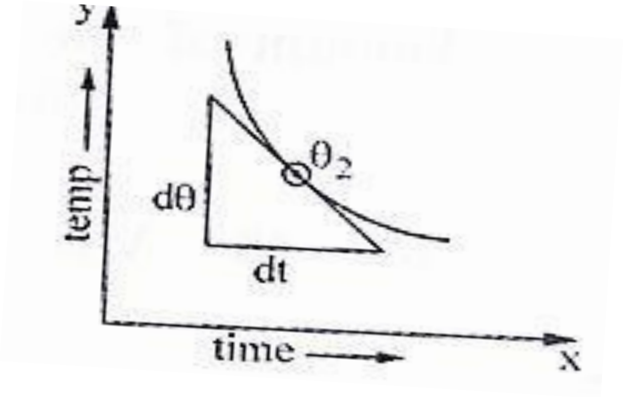
$$\frac{K\ \pi r^2\ (\theta_1 - \theta_2)}{t} = m\ S\ Rc$$

$$K = \frac{\mathrm{m\ S\ t\ Rc}}{\pi r^2\ (\theta_1 - \theta_2)} \quad \dashrightarrow (3)$$



To find the rate of cooling (R_c)

In eqn. (3) R_c represents the rate of cooling of the slab along with the steam chamber. To find the rate of cooling for the slab alone, the bad conductor is removed and the steam chamber is directly placed over the slab and heated



When the temperature of the slab attains 5°C higher than θ_2, the steam chamber is removed. The slab is allowed to cool, as shown in the fig, simultaneously starting stop watch.

A graph is plotted taking time along ‘x’ axis and temperature along ‘y’ axis, the rate of cooling for the slab alone. i.e., $\frac{d\theta}{dt}$ is found from the graph as shown in the fig. The rate of cooling is directly proportional to the surface area exposed

Case (i)

Steam chamber and bad conductor are placed over slab, in which radiation takes place from the bottom surface of the area (πr^2) of the slab and the sides of the slab of area ($2 \pi r h$).

$$\therefore R_c = \pi r^2 + 2\pi r h$$

$$R_c = \pi r (r + 2 h) \quad \longrightarrow (4)$$

Case (ii)

The heat is radiated by the slab alone, i.e., from the bottom of area (πr^2), top surface of the slab of area πr^2 and also through the sides of the slab area $2 \pi r h$.

$$\left(\frac{d\theta}{dt}\right)_{\theta_2} = \pi r^2 + \pi r^2 + 2\,\pi\,\mathrm{r\,h}$$

$$\left(\frac{d\theta}{dt}\right)_{\theta_2} = 2\,\pi\,\mathrm{r\,(r + h)} \dashrightarrow (5)$$

From eqn (4) and (5)

$$\frac{R_c}{\left(\frac{d\theta}{dt}\right)_{\theta_2}} = \frac{\pi\,\mathrm{r\,(\,r + 2\,h)}}{2\,\pi\,\mathrm{r\,(r + h)}}$$

$$R_c = \frac{\mathrm{(\,r + 2\,h)}}{2\,\mathrm{(r + h)}}\left(\frac{d\theta}{dt}\right)_{\theta_2} \dashrightarrow (6)$$

Substituting eqn. (6) in eqn. (3),

$$K = \frac{m\,s\,t\,\left(\frac{d\theta}{dt}\right)_{\theta_2}\,(\,r + 2h)}{\pi r^2(\theta_1 - \theta_2)\,2\,(r + h\,)}\,W\,m^{-1}K^{-1} \qquad \dashrightarrow (6)$$

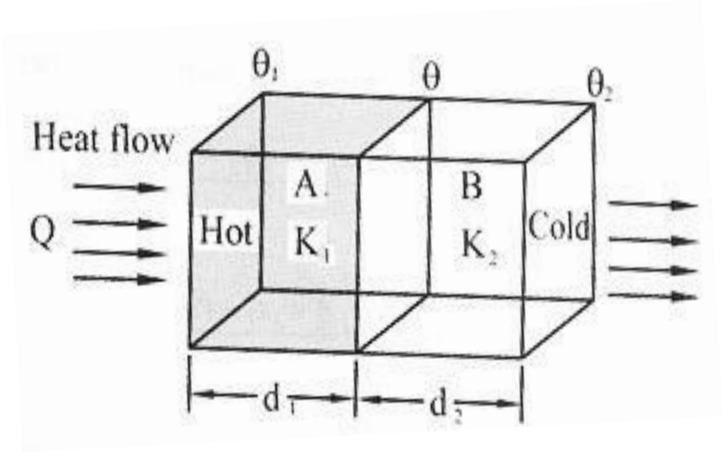
Hence the thermal conductivity of the given bad conductor can be determined from the above formula.

4.12. HEAT CONDUCTION THROUGH A COMPOUND MEDIA (SERIES)

i) Bodies in series

Consider a composite slab or compound wall of two different (heterogeneous) materials, A and B of thermal conductivity K_1 and K_2 respectively.

Let the thickness of these two layers A and B be d_1 and d_2 respectively as shown in fig. Let the temperature of the end faces be θ_1 and θ_2 and temperature at the conduct surface be θ, which is unknown. Heat will flow from face-A to face-B through surface of contact only if $\theta_1>\theta_2$. After steady state is reached heat flowing per second (Q) through every layer is same.



Amount of heat flowing per second through A

$$Q = \frac{K_1 A\,(\theta_1 - \theta)}{d_1} \qquad \dashrightarrow (1)$$

Amount of heat flowing per second through B

$$Q = \frac{K_2 A\,(\theta - \theta_2)}{d_2} \qquad \dashrightarrow (2)$$

Where A is Area of cross-section of both layers, which is same. Since the amount of heat flowing through A and B are equal, we can write, Equation (1) = Equation (2)

$$\frac{K_1 A\,(\theta_1 - \theta)}{d_1} = \frac{K_2 A\,(\theta - \theta_2)}{d_2}$$

$$\frac{K_1\theta_1}{d_1} - \frac{K_1\theta}{d_1} = \frac{K_2\theta}{d_2} - \frac{K_2\theta_2}{d_2}$$

Rearranging we get $\frac{K_1\theta_1}{d_1} + \frac{K_2\theta_2}{d_2} = \left(\frac{K_1}{d_1} + \frac{K_2}{d_2}\right)\theta$

$$\theta = \frac{\frac{K_1\theta_1}{d_1} + \frac{K_2\theta_2}{d_2}}{\frac{K_1}{d_1} + \frac{K_2}{d_2}} \qquad --\rightarrow (3)$$

Equation (3) gives us the expression for temperature at the interface of the compound media which is having two layers. By substituting the value θ in equation (1), we get

$$Q = \frac{K_1 A}{d_1}\left[\theta_1 - \left(\frac{\frac{K_1\theta_1}{d_1} + \frac{K_2\theta_2}{d_2}}{\frac{K_1}{d_1} + \frac{K_2}{d_2}}\right)\right]$$

$$Q = \frac{K_1 A}{d_1}\left(\theta_1 - \left(\frac{(d_2K_1\theta_1 + d_1K_2\theta_2)/d_1d_2}{(K_1d_2 + K_2d_1)/d_1d_2}\right)\right)$$

$$Q = \frac{K_1 A}{d_1}\left(\frac{(d_2K_1 + d_1K_2)\theta - d_2K_1\theta_1 - d_1K_2\theta_2}{(K_1d_2 + K_2d_1)}\right)$$

$$Q = \frac{\mathrm{K_1A}}{\mathrm{d_1}}\left(\frac{(\mathrm{d_2K_1}\theta_1 + \mathrm{d_1K_2}\theta_1) - \mathrm{d_2K_1}\theta_1 - \mathrm{d_1K_2}\theta_2}{(\mathrm{K_1d_2} + \mathrm{K_2d_1})}\right)$$

$$\mathrm{Q} = \frac{\mathrm{K_1A[K_2d_1}(\theta_1 - \theta_2)}{\mathrm{d_1(K_1d_2 + K_2d_1)}}$$

$$Q = \frac{\mathrm{K_1K_2A}(\theta_1 - \theta_2)}{(\mathrm{K_1d_2 + K_2d_1})} = \frac{\mathrm{A}(\theta_1 - \theta_2)}{\frac{\mathrm{K_1d_2}}{\mathrm{K_1K_2}} + \frac{\mathrm{K_2d_1}}{\mathrm{K_1K_2}}}$$

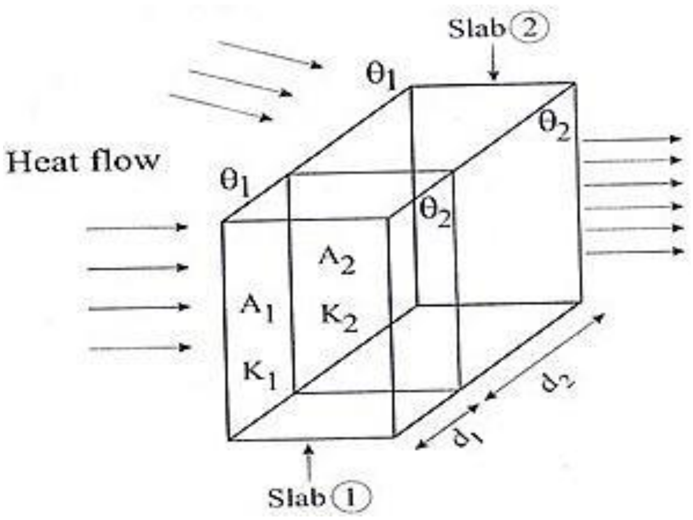
$$\mathrm{Q} = \frac{\mathrm{A}(\theta_1 - \theta_2)}{\frac{\mathrm{d_1}}{\mathrm{K_1}} + \frac{\mathrm{d_2}}{\mathrm{K_2}}} \qquad --\rightarrow (4)$$

Equation (4) gives the amount of heat conducted by the two layers in series.

4.13 HEAT CONDUCTION THROUGH A COMPOUND MEDIA (PARALLEL)

ii) Bodies in parallel

Consider a composite slab or compound wall of two different (heterogeneous) materials A and B of thermal conductivities K_1 and K_2 and of thickness d_1 and d_2 respectively. These two material layers are arranged in parallel as shown in fig. The opposite faces of the material are kept at temperatures θ_1 and θ_2 respectively.



Let A_1 and A_2 be the areas of cross-sections of the materials. The amount of heat flowing through the first slab

$$Q_1 = \frac{K_1 A_1 (\theta_1 - \theta_2)}{d_1} \quad --\rightarrow (1)$$

The amount of heat flowing through the second slab

$$Q_2 = \frac{K_2 A_2 (\theta_1 - \theta_2)}{d_2} --\rightarrow (2)$$

The total heat flowing through these two slabs per second

$$Q = Q_1 + Q_2 \dashrightarrow (3)$$

Substituting eqn. (1) & (2) in eqn. (3), we get,

$$Q = \frac{K_1A_1}{d_1}(\theta_1 - \theta_2) + \frac{K_2A_2}{d_2}(\theta_1 - \theta_2)$$

$$Q = \left[\frac{K_1A_1}{d_1} + \frac{K_2A_2}{d_2}\right](\theta_1 - \theta_2) \quad \dashrightarrow (4)$$

Equation (4) gives us the amount of heat flowing through compound wall of two layers in parallel.

4.14 FORMATION OF ICE ON PONDS

Consider a layer of ice x cm thick on the surface of a pond. Let the temperature of air above the surface of ice be -θ° C and that of water below ice be 0° C. Suppose a thickness 'dx' of ice is formed in time 'dt'. Then,

Mass of ice formed= $Adx\rho$

Here

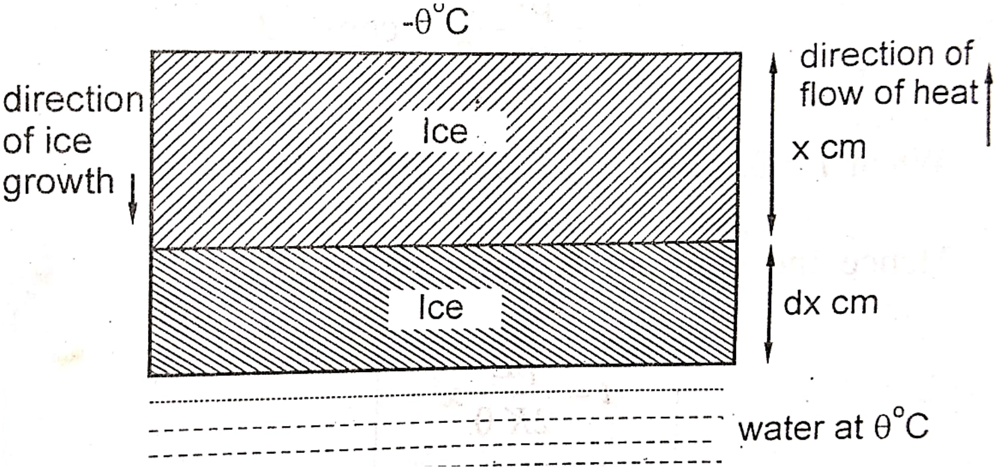
A= Area of the pond

ρ= the density of ice

L= the latent heat of fusion of ice

Heat lost by water= AdxρL calories

This heat is conducted across a layer of ice of thickness 'x' upwards.



Heat conducted= $KA\frac{\theta}{x}dt$calories

$$KA\frac{\theta}{x}dt = \text{AdxρL}$$

$$\frac{dx}{dt} = \frac{K\theta}{\rho Lx}$$

$\frac{dx}{dt}$- Represents the rate of growth of the thickness of ice

$$dt = \left(\frac{\rho L}{K\theta}\right)xdx$$

Total time taken by the layer of ice to increase in thickness by ‘x’

$$\int dt = \frac{\rho L}{K\theta}\int xdx$$

$$t = \frac{\rho L}{K\theta}\frac{x^2}{2} + constant$$

When t=0, x= 0

Hence the constant is zero

$$t = \frac{\rho L}{2K\theta} x^2$$

4.15 THERMAL INSULATION

Thermal insulation is defined as blocking of reducing heat transfer between two objects. It is also defined as thermal resistance.It also provides a region of insulations in which thermal conduction is reduced. In other way we can say that the thermal radiation is reflected rather than absorbed by the body at lower temperature.The insulating capacity of the material is mainly based on the thermal conductivity, insulating capacity, density and specific heat capacity of the material. Thermal insulation mainly observed in following areas.

(i).Buildings (ii) Mechanical systems (iii)Space Craft (iv) Automotive clothing etc.The following factors are mainly to be considering for providing good thermal effect, viz.(i)Thermal insulation (ii) Thermal comfort (iii) Thermal regulation (iv) Thermal resistivity

4.15.1. CLASSIFICATION OF THERMAL INSULATING MATERIALS

There are three types of thermal insulating materials, viz,

(i). **Fibrous thermal insulating materials** such as felt, fur, wool and hair.

(ii). **Granular thermal insulating materials** such sand and sawdust.

(iii). **Porous thermal insulating materials** such as asbestos, aluminium foil, form rubber, cork, expended ebonite etc, all of which have a number of tiny air-pockets within them.

The material can also be classified into two categories based on their density.

S.No.	Less density insulating materials	High density insulating materials
1	It has high resistive power	It has low resistive power
2	The materials can be quickly heated and cooled	The material take more time both heating and cooling
3	These materials will not absorb more heat	These materials will absorb enormous heat
4	Example: cork, cardboard thermo cool etc.	Examples: Brick, stone concrete etc.

4.15.2. PROPERTIES OF THERMAL INSULATING MATERIALS.

1. The material should be fire proof.
2. It should have high volumetric specific heat.
3. It should have low thermal conductivity.
4. It should be a poor absorber of moisture.
5. It should withstand for any environmental conditions.
6. It should have a fibrous granular porous structure.
7. It should be of low cost.
8. It should good looking.

4.16. HEAT EXCHANGERS

Definition

A heat exchanger is a device that is used to transfer the heat between a solid and liquid (or) between two (or) more liquid, without mixing. It is used to reduce the heat produced by a device (or) machine.

When a fluid is used to transfer heat, the fluid could be a liquid, such as water or oil, or could be moving air. Heat exchangers are used to transfer heat from one medium to another. These media may be a gas, liquid, or a combination of both. The media may be separated by a solid wall to prevent mixing or may be indirect contact. Heat exchangers can improve a system's energy efficiency by transferring heat from systems where it is not needed to other systems where it can be usefully used.

Measurement:

In a heat exchanger, the driving temperature across the heat transfer surface varies with position. Therefore the temperature difference is measured only in terms of log mean temperature difference (LMTD)

Mode of heat flow

The basic component of a heat exchanger can be viewed as a tube with one fluid running through it and another fluid flowing by on the outside. There are thus three heat transfer operations that need to be described:

1. Convective heat transfer from fluid to the inner wall of the tube,
2. Conductive heat transfer through the tube wall, and
3. Convective heat transfer from the outer tube wall to the outside fluid.

Heat exchangers are typically classified according to flow arrangement and type of construction. The simplest heat exchanger is one for which the hot and cold fluids move in the same or opposite directions in a concentric tube (or double-pipe) construction. In the parallel-flow arrangement of Figures 4.1, the hot and cold fluids enter at the same end, flow in the same direction, and leave at the same end. In the counter flow arrangement of Figures 4.2 There are 3 modes by which heat is transferred from one fluid to another in all the heat exchanges, viz (i).Parallel- flow heat exchanger

(ii).Counter –flow heat exchanger

(iii). Cross –flow heat exchanger

4.16. 1. Parallel- flow heat exchanger

In parallel flow heat exchanger, the two fluids (one cold & other hot) enter the exchanger at the same end and travel in parallel to one another as shown in figure and reach the other end, and thereby exchange the heat.

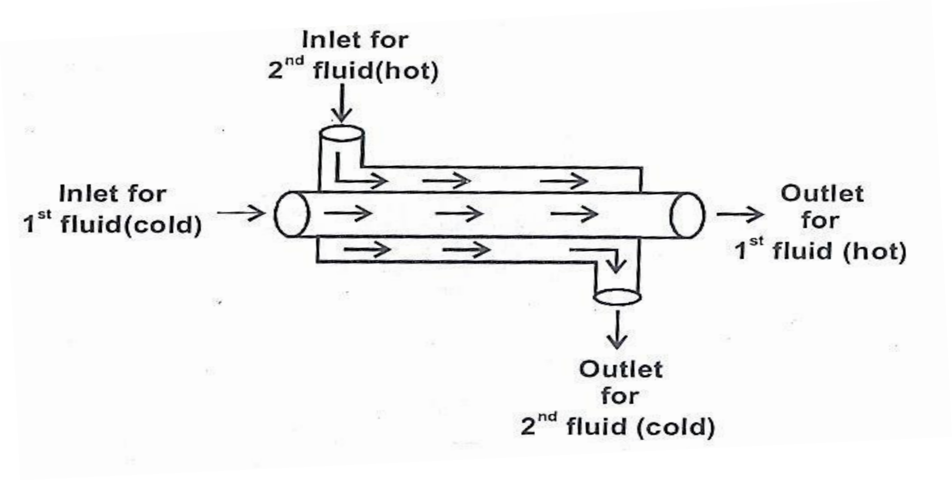


Fig. 4.1

In this type the temperature difference between the two fluids is large at the entrance end and becomes very small at the exit end.

4.16. 2.Counter –flow heat exchanger

In Counter flow heat exchanger, the two fluids (one cold & other hot) enter the exchanger at the opposite ends as shown in figure and exchange the heat between each other.

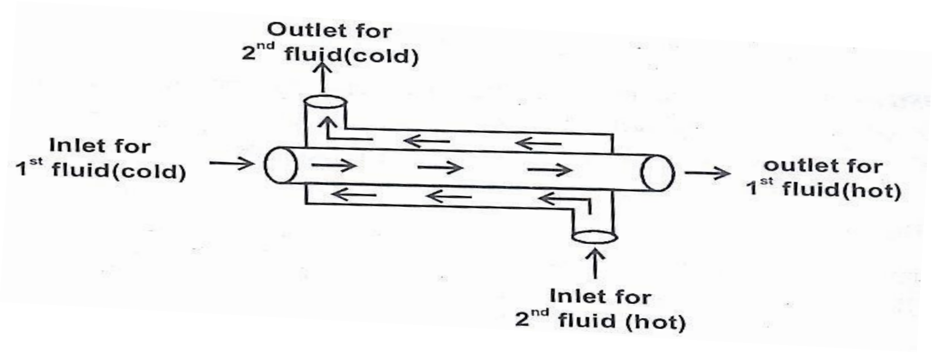
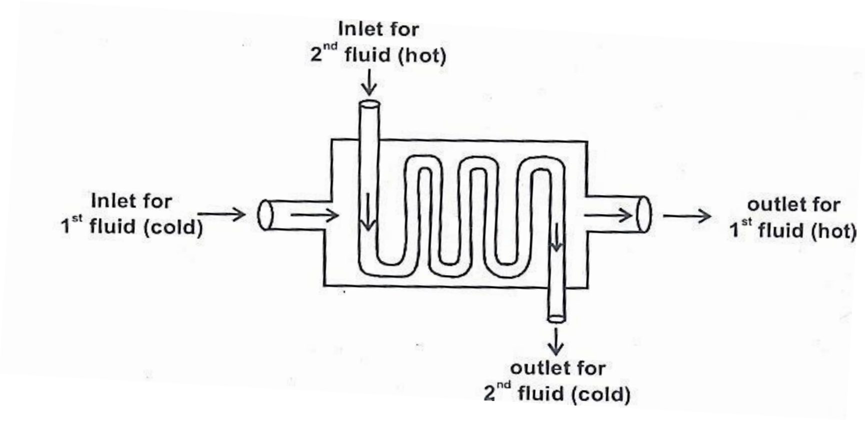


Fig. 4.2

It is the most commonly used heat exchanger for liquid- liquid heat transfer and is more efficient, when compared to parallel flow heat exchanger.

4.16. 3. Cross –flow heat exchanger



In cross flow heat exchanger, the fluids (one cold & other hot) travel roughly perpendicular to each other, through the heat exchanger as shown in figure and exchanges the heat between them.

In this case the heat transfer take place between a liquid flowing inside a tube (or) tubes (many coils can be designed) it is mainly used in solar water heater, car radiators, refrigerators and Air conditioners.

4.16. 4. Types

There are different types of heat exchangers, which are used in engineering and industries field viz.

1. Fluid heat exchanger.
2. Shell and tube heat exchanger.
3. Plate heat exchanger.
4. Plate and shell heat exchanger
5. Plate fin heat exchanger.
6. Pillow plate heat exchanger.
7. Direct contact heat exchanger.
8. Micro- channel heat exchanger etc.

4.16.5 Regenerative heat exchangers

In a regenerative heat exchanger, the flow path normally consists of a matrix, which is heated when the hot fluid passes through it (this is known as the "hot blow"). This heat is then released to the cold fluid when this flows through the matrix (the "cold blow"). Regenerative Heat Exchangers are sometimes known as ***Capacitive Heat Exchangers***. Regenerators are mainly used in gas/gas heat recovery applications in power stations and other energy intensive industries. The two main types of regenerator are Static and Dynamic. Both types of regenerator are transient in operation and unless great care is taken in their design there is normally cross contamination of the hot and cold streams. However, the use of regenerators is likely to increase in the future as attempts are made to improve energy efficiency and recover lower grade heat. However, because regenerative heat exchangers tend to be used for specialist applications recuperative heat exchangers are more common.

Recuperative heat exchangers

There are many types of recuperative exchangers, which can broadly be grouped into indirect contact, direct contact and specials. Indirect contact heat exchangers keep the fluids exchanging heat separate by the use of tubes or plates etc.. Direct contact exchangers do not separate the fluids exchanging heat and in fact rely on the fluids being in close contact.

Overall heat transfer coefficient

The overall heat transfer coefficient is employed in calculating the rate of heat transfer Q from one fluid at an average bulk temperature T_1 through a solid surface to a second fluid at an average bulk temperature T_2 (where $T_1 > T_2$). The defining equation is generally only applicable to an incremental element of heat transfer surface dA for which the heat transfer rate is dQ , and the equation is strictly valid only at steady state conditions and negligible lateral heat transfer in the solid surface, conditions generally true enough in most practical applications. The defining equation is

$$d\dot{Q} = U(T_1 - T_2)dA$$

where U is referenced to a specific surface (see below).

In the particular situation of heat transfer across a plane wall of uniform thickness, U is related to the individual film heat transfer coefficients, α_1 and α_2, of the two fluids by the equation

$$U = \frac{1}{\frac{1}{\alpha_1} + \frac{\delta_w}{\lambda_w} + \frac{1}{\alpha_2}}$$

where δ_w is the thickness of the wall and λ_w is the thermal conductivity of the wall.

If there are fouling deposits on the wall, they have a resistance to heat transfer, R_1 and R_2, in units of m^2 K/W, and these resistances must be added in

$$U = \frac{1}{\frac{1}{\alpha_1} + R_1 + \frac{\delta_w}{\lambda_w} + R_2 + \frac{1}{\alpha_2}}$$

For the special but very important case of heat transfer through the wall of a plain round tube, the different heat transfer areas on the inside and outside surfaces of the tube need to be considered. Let dA_i be the inside incremental area and dA_o be the outside. Then (including fouling resistances R_{fi} and R_{fo} inside and out):

$$U = \frac{1}{\frac{1}{\alpha_1} + R_{fi} + \frac{r_i \ln(r_0 / r_i)}{\lambda_w} + R_{fo} + \frac{r_i}{r_0} + \frac{r_i}{\alpha_0 r_0}}$$

where U_i is termed the "overall heat transfer coefficient referenced to (or based on) the inside tube heat transfer area", and r_i and r_o the inside and outside radii of the tube.

Alternatively, the overall coefficient may be based on the outside heat transfer area, giving

$$U = \frac{1}{\frac{r_0}{\alpha_i f_i} + R_{fi} + \frac{r}{r_i} + \frac{r_0 \ln(r_0 / r_i)}{\lambda_w} + R_{f0} + \frac{1}{\alpha_0}}$$

where U_o is termed the **overall heat transfer coefficient** based on the outside tube heat transfer area

4.16.6 Applications

Heat exchangers have a very wide range of applications some of them are detailed below.

- They are used in refrigerators, air conditioners etc,.
- Heat exchangers are often used in power plants (or) engines to cool the exhaust hot gases.
- They are widely used in petroleum refineries , petro-chemical plants etc.,
- They are also used in natural gas processing and sewage water treatment plants.
- Heat exchanges are found in internal combustion engines, radiator coils, etc.,
- Use less water especially hot water. Also use water treatment.

4.17 REFRIGERATORS

Generally heat cannot flow from cold body to hot body. It is possible to do so, if some external work (or) pressure is done the working substance. This concept is used in refrigerators.

General terminologies

1. Refrigeration

Refrigeration is process of reducing and maintains the temperature of a body below the normal temperature or it is the process of removing the heat from a substance under controlled conditions.

2. Refrigerator

Refrigerator is an equipment used to reduce and maintaining the temperature of the body below atmospheric temperature. It is obtained by removing the heat from the space continuously.

3. Refrigerant

Refrigerant is a fluid which absorbs the heat from the body and rejects the heat at high temperature.

Examples: ammonia, carbon-di-oxide, Freon, methyl chloride, chloro Fluro carbins (CFC).

4. Capacity of Refrigerator (or) Refrigerating effect.

It is the amount of heat extracted from the cold body per unit mass per second. It is also defined as the rate at which refrigeration produced is called the capacity of refrigerator. It is expressed in 237 tonne of refrigeration.

5. Tonnes of refrigeration

A 238 tonne of refrigeration is defined as the amount of refrigeration effect produced by uniform melting of one 238 tonne (1000Kg) of ice at 0^0C water in 24 hours.

1tonne of refrigeration = 210 KJ/ min (or) 3.5 KJ/ Sec.

6. Performance coefficient

Coefficient of performance (COP) is the ratio of heat extracted and work done

$$\text{COP} = \frac{\text{Heat Extracted}}{\text{Work done}}$$

Type of refrigerators

There are two types of refrigerator

- Vapour Compression refrigerator -(Example: Domestic refrigerators)
- Vapour absorption refrigerators- (Example: Commercial refrigerators)

Let as discuss about s domestic refrigerator.

4.17.1 Domestic Refrigerator

Principle

The second law of the thermodynamics as given by clausius, “without doing external work it is impossible to transfer heat from a cold body to a hot body” is the principle of refrigerator.

Here the ammonia (working Substance) takes heat from the refrigerator and due to external work done on ammonia, it give heat to atmospheric air (at normal (or) high temperature) and keep the refrigeration continuously cool.

Design

It consists of two coils viz.

1. Evaporator coil to convert liquid ammonia to vapour and
2. Condenser coil to convert vapour to liquid ammonia, as shown in figure.

The compressor in the refrigerator is used to compress the ammonia vapour using a piston to a very high pressure and helps in doing the external work on the ammonia. The whole setup is kept in well air circulation area for better performance.

Working

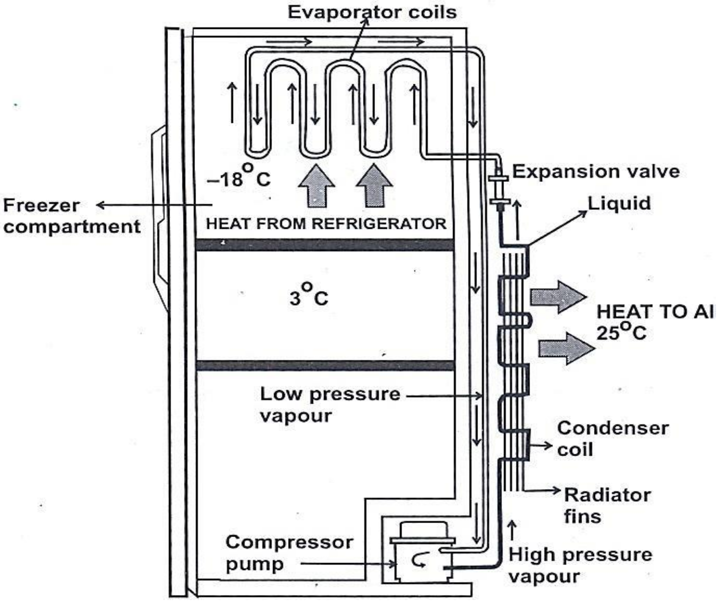
1. In domestic refrigerator liquid ammonia is used as a working substance for

cooling the refrigerator

2. Here liquid ammonia at low pressure is passed the evaporator coils, wherein it expends and absorbs the heat from the refrigerator.

3. The liquid ammonia takes up the heat from the refrigerator and is converted into low pressure vapour.

4. Now the compressor is used to compress the ammonia vapour externally using a piston, to very high pressure.



5. This ammonia at high pressure is allowed to pass through for condenser coils.

6. While passing the ammonia vapour gives heat to the atmosphere air at room temperature and become liquid ammonia again due to cooling.

7. The cool liquid ammonia in turn acts as a primary refrigerant and keeps the refrigerator cool.

8. This cycle of processes continues and makes the refrigerator to be in cool condition always.

Applications

1. It is used for preserving the food, fruits and drinks for a long duration.
2. It is used to preserve flowers medicines and medical durgs.
3. Refrigerator is used to manufacture ice plants.
4. In industries they are used for processing lubricants, rubber, steel etc.,
5. It is used for producing frozen foods, ice creams, chemical and other products.

Advantages

1. Cost of refrigerant is low
2. It used to store the food for a long time.
3. It protects the food from direct sun light.
4. It products the food from microbes, insects and rodents.

Disadvantages

1. It consumes large amount of electricity
2. It causes global warming
3. Harmful pollutant gas like CFC (Choloroflurо carbon) is used in refrigeration.
4. Preserving food in refrigerator not good for health.

4.18. OVENS

Ovens is a cylindrical shape material made up of heavy cast-iron, which is used to roast or cook food in ancient days. In modern days ovens have become more hightech, in which electricity is used as the source for the heating coils. More recently micro-wave oven become very popular for cooking food.

Types of Ovens

There are different types of ovens, based on their usage and their development from earlier Greek age to modern era. They are as follows'

1. Earth oven

An earth oven is the pit dug into the found and food wrapped in leaves (or) animal skin is heated wood (or) rocks (or) debris.

2. Ceramic oven

The ceramic oven in constructed using clay (or) some ceramic materials with bricks and the food was cooked using firewood. It is still persisting in India and referred as tandoor.

3. Gas oven

In gas ovens the food was prepared either on a gas stove (or) a oven compartment in gas stoves. Usually thermostats are used in gas ovens, which are used to regulate the temperature, while cooking food.

4. Toaster oven

Toaster ovens are small electric ovens with a front door, wire rack and removable baking pan. Here bread is placed horizontally (or) vertically on the racks and was toasted by using electricity.

5. Microwave oven.

Microwave oven are the modern ovens which uses tiny, high powered microwaves, for cooking the food. Here the microwave directly heat the molecules in the food particles similar to that of how sun heats our face by radiation. Let us discuss in detail about micro-wave oven.

4.19 SOLAR POWER

Solar power is the processes of converting (or) utilizing the abundantly available solar energy either directly as heat (or) indirectly by converting it into electrical power using photo-voltaic cell.

Estimation of solar power

We have known that the sun is the powerful fusion reactor, in which hydrogen atoms fuse to form helium and in turn releases tremendous amount of energy.

Due to nuclear fusion the temperature of the sun at the core region is 15 million degree Celsius. Whereas, the surface temperature in the photosphere of the sun 6000K.

This solar energy reaches the earth at the rate of 1.4 KW per square meter of the surface perpendicular to the direction of the sun rays. The mean radius of the earth is $1.5 \times 10^{11}\ m$

The total power radiated by the sun can be written as

$$\rho = Energy\ received\ by\ the\ earth\ \times surface\ area\ of\ the\ earth$$

$$\rho = Energy\ received\ by\ the\ earth\ \times 4\pi r^2$$

$$\rho = 1.4 \times 10^3\ \times 4\pi r^2 = 1.4 \times 10^3\ \times 4 \times 3.14 \times (1.5 \times 10^{11})^2$$

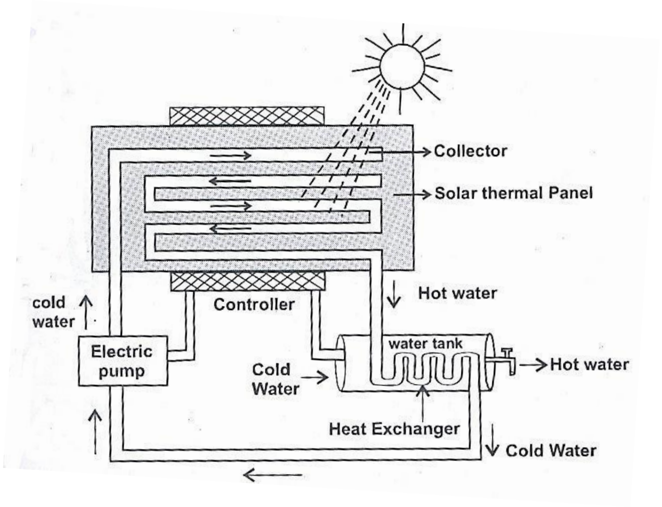
$$\rho = 3.956 \times 10^{26}$$

4.19.1. Solar Water heater

Solar water heater is based on the principle of converting solar energy into electrical energy and then into heat energy, using solar electric panels, so called solar cells (or) photo voltaic cells. Nowadays, solar thermal panels were widely used, which converts the solar energy directly into heat energy.

Construction

1. A simple roof top solar water heater is as shown in figure.
2. It consists of a solar thermal panel in which collector is used to collect, capture and retain the heat radiations from the sun.
3. The heat exchanger is made up of coil of copper pipes and is kept inside the water tank.
4. Heat exchanger is used to transfer that heat energy from the hot water passing through the copper pipe in the heat exchanger to the cold water in the water tank.



5. An electric pump is used to pump the cold water coming out from the heat exchanger to the collector of the thermal panel.

6. The total system is controlled by the controller unit, which is used to

    i) Fill the water with auto cut-off in the tank.

    ii) Switch on / off the electric pump, whenever required.

Working

- Initially, with the help of the Electric pump, cold water is pumped to the collector in the solar thermal panel.
- Now due to thermal radiations that fall on the thermal panel, the water in collector is heated up.
- This hot water is allowed to pass through the water tank with the help of the heat exchangers.
- The heat exchanger, which is made up of coil of copper pipes, transfers the heat energy from the water inside the copper coils, to the water presents in the water tank and therefore the water in the water tank becomes ho.
- The water coming out from the heat exchanger, after transferring the energy, become cold water and enters into the electric pump again as shown in figure.
- Thus we can run off hot water from the tank at any time without affecting the panels operation.

Advantages

1. Solar energy is free and abundant.

2. A solar thermal panel occupies less space compeered to solar photovoltaic panels.

3. Solar thermal panels are 3 times more efficient, approximately 80% of the thermal radiation is turned into heat energy.

4. We can save money by paying less electricity bill.

5. It an Eco friendly way to heat water for the domestic need.

Disadvantages

1. Capital investment and installation cost is high.

2. Annual maintenance is required to check the working of pump, anti- freezing etc.

3. It occupies space and depends on the availability of direct sun light.

4. It is not useful during rainy (or) foggy days.

5. It will take about 10 to 15 years to get back the money we spent for capital investment and installation, which is based on your usage.

Thermal Convection

Convection is a process of heat transfer due to movement of molecules within the fluids. Convection above a hot surface occurs because hot air expands, becomes less dense, and rises (see Ideal Gas Law). Hot water is likewise less dense than cold water and rises, causing convection currents which transport energy. Convection can also lead to circulation in a liquid, as in the heating of a pot of water over a flame. Heated water expands and becomes more buoyant.

Radiative Properties

- When radiation strikes a surface, a portion of it is reflected, and the rest enters the surface.

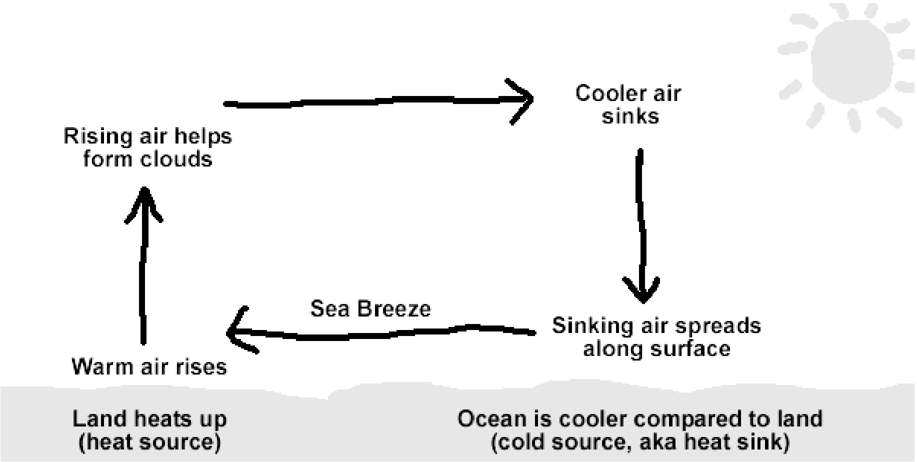
- Of the portion that enters the surface, some are absorbed by the material, and the remaining radiation is transmitted through.

- The ratio of reflected energy to the incident energy is called *reflectivity*, ρ.
- Transmissivity (τ) is defined as the fraction of the incident energy that is transmitted through the object.

4.20. SEA BREEZES AND LAND BREEZES

Sea breezes are winds that are caused by the differential heating rates of land and sea. During the day the land heats up more rapidly than the sea because:

- land has a lower specific heat than water – it requires less heat to raise the temperature of 1 kilogram of earth by 1 degree than it does for water;
- the incoming heat warms a shallower layer of the land because earth is a poor conductor of heat. In water the heat is spread through a significant depth by conduction, convection and focused mixing.
- Sea breezes tend to be stronger in the tropics where greater temperature differences are established between land and water and there is also a greater tendancy for instability over the heated land.
- Lake breezes are also formed around large areas of inland water.
- Low level wind shears are often enhanced by sea breeze circulation.
- Land breezes because the of reverse situation. Overnight the land cools to be colder than the sea so cold overland air moves to replace warmer the warmer air over the sea.

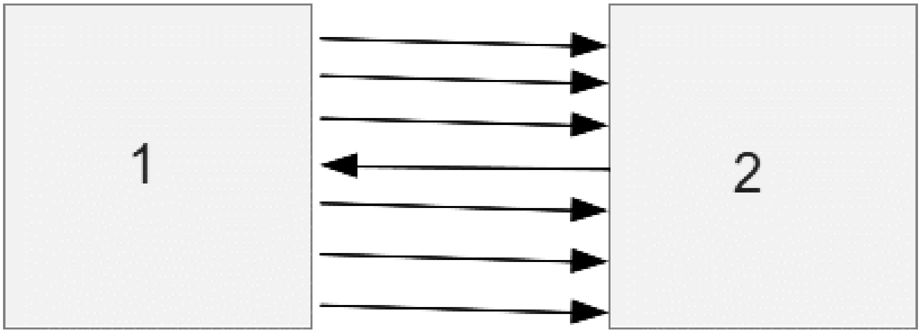


A land breeze is created when the land is cooler than the water such as at night and the surface winds have to be very light. When this happens the air over the water slowly begins to rise, as the air begins to rise the air over the surface of the ocean has to be replaced, this is done by drawing the air from the land over the water, thus creating a sea breeze. Usually land breezes are light and reach only 5 to 8 knots.

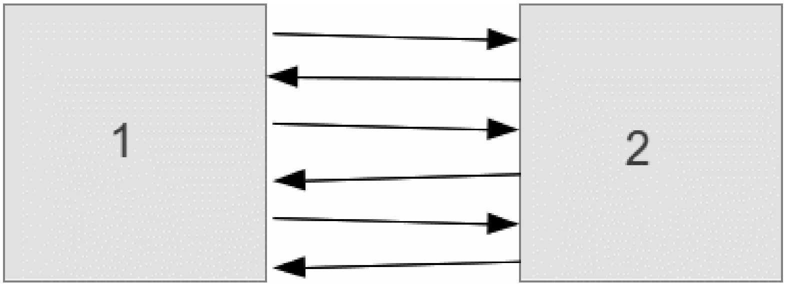
A sea breeze is created when the surface of the land is heated sufficiently to start air rising. As air rises, it is replaced by air from the sea; you have now created a sea breeze. Sea breezes tend to be much stronger and can produce gusty winds as the sun can heat the land to very warm temperatures, thereby creating a significant temperature contrast to the water. Speeds of 10 to 20 knots is not uncommon for a sea breeze.

4.21 PREVOST THEORY OF HEAT EXCHANGE

Prevost suggested that all bodies radiate energy, but hotter bodies radiate more heat than colder bodies. Suppose for example that we have two bodies 1 and 2 at different temperatures T_1 and T_2 with T1 is greater than T_2. Each body will then radiate heat to the other but since $T_1 > T_2$ the energy radiated from 1 to 2 is greater than the energy radiated from 2 to 1.



Heat will be transferred from body 1 to body 2 as long as $T_1 > T_2$. Eventually the bodies will be at the same temperature so $T_1 = T_2$and heat will be transferred from body 1 to 2 at the same rate as from 2 to 1. The bodies will then be in thermal equilibrium.



According to Prevost, a body emits heat radiations at all temperatures which is independent of surroundings but the quantity of heat radiated per second depends on the temperature of the body. The rate of radiation is greater at a higher temperature than at a lower one. At the same time, the body absorbs thermal radiation from the surrounding. If the rate of emission by the body is less than the rate of absorption from surroundings, there is a net heat gained by the body and its temperature rises. Conversely, if the body is emitting at the rate higher than the absorption from the surroundings, there is a net loss of heat, therefore, there is a fall of temperature of the body. The rate of emission will be greater if the temperature of the body is higher than that of the surrounding and vice-versa.

4.22 HEAT RADIATION

Radiation is a method of heat transfer that does not rely upon any contact between the heat source and the heated object. Both conduction and convection require matter to transfer heat. Heat can be transferred from one place to another by three methods: conduction in solids, convection of fluids (liquids or gases), and radiation through anything that will allow radiation to pass. The method used to transfer heat is usually the one that is the most efficient. If there is a temperature difference in a system, heat will always move from higher to lower temperatures.

4.22.1 Heat Emission

Heat emission from a solid body into liquid or gas at their convective agitation is connected with thermal conductivity through the thin contact layer of the static matter adjoining at the body surface. As the temperature gradient is proportional to temperature difference of the surface temperature and the bulk temperature of fluid or gas, the equation of heat emission has the form of Newton's heat emission law:

$$\frac{dT}{dt} = \alpha(T - T_s)s$$

where T is the body temperature, T_s is the temperature of liquid or gas, α is the heat emission coefficient, and s is the area of the surface

4.22.2 Heat absorption

Heat absorption refers to the heat transfer that occurs between two bodies; it can occur through conduction, convection or radiation. Heat absorption also is an endothermic reaction. In an endothermic process, a cooler object absorbs the hotter object's heat.

4.22.3 Thermal emissivity

Emissive power is the energy of thermal radiation emitted in all directions per unit timeper unit area of a surface at any given temperature.

4.23 BLACK BODY RADIATION

A blackbody is an idealized object which absorbs and emits all frequencies. Classical physics can be used to derive an equation which describes the intensity of blackbody radiation as a function of frequency for a fixed temperature — the result is known as the Rayleigh-Jeans law. Although the Rayleigh-Jeans law works for low frequencies, it diverges as f2; this divergence for high frequencies is called the ultraviolet catastrophe.

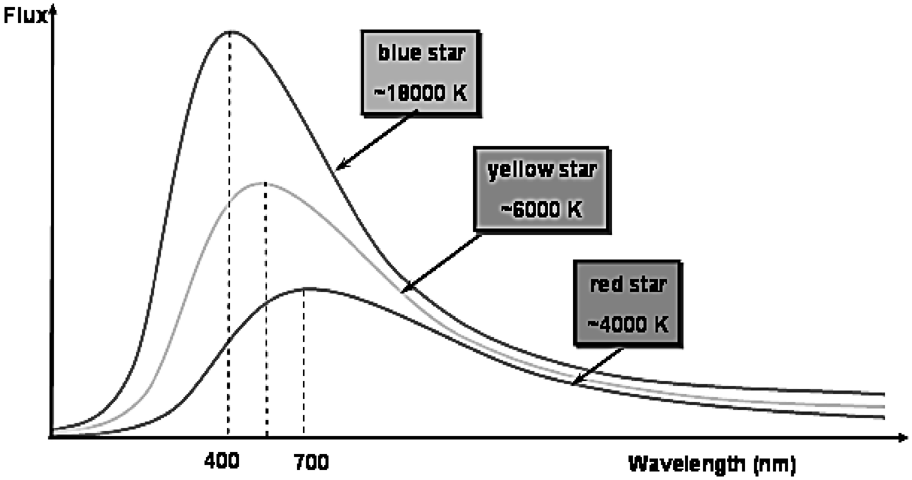
A perfect black body is one which absorbs and emits in all the radiations (corresponding to all wavelengths) that fall on it. The radiation given out by a perfect black body is called Black body radiation.

In practice a perfect black body is not available. Therefore let us consider a hollow sphere coated with lamp black on its inner surface. Now when the black body is placed in a temperature bath of fixed temperature, the heat radiations will come out only through the hole in the sphere and not through the walls of the sphere. Therefore, we can conclude that the radiations are emitted from the inner surface of the sphere and not from the outer surface of the sphere. Thus a perfect black body is a perfect absorber and also a perfect radiator of all wavelengths. In simple terms a blackbody is a volume which operates as a perfect absorber and emitter of thermal radiation. More specifically a blackbody is a volume which:

1. Absorbs all incident radiation independent of wavelength and direction.
2. Will emit more energy than any other surface at a given temperature.
3. Emits radiation independent of direction, i.e. a diffuse emitter.

4.22.1 Energy spectrum

When a perfect black body is allowed to emit radiations at different temperatures, then the distribution of the energy for different wavelengths at various temperatures is obtained as shown in the figure.



From figure the following results are formulated.

i. The energy distribution is not uniform for a given temperature.

ii. The intensity of radiation (E) increases with respect to the increase in wavelength at particular wavelength in becomes maximum (λm) and after this it starts decreasing with respect to the increase in wavelength.

iii. When the temperature is increased, the maximum wavelength (λm) decreases.

iv. For all the wavelengths an increase in its temperature causes increase in energy.

iv. The total energy emitted at any particular temperature can be calculated from the area under that particular curve.

4.24 KIRCHHOFF’S LAW

Consider an enclosure as shown in fig.7.2 and a body is placed inside the enclosure. The radiant heat flux (q) is incident onto the body and allowed to come into temperature equilibrium. The rate of energy absorbed at equilibrium by the body must be equal to the energy emitted.

$$EA = \alpha qA$$

$$E = \alpha q \qquad (1)$$

where, E is the emissive power of the body, α is absorptivity of the of the body at equilibrium temperature, and *A* is the area of the body.

Now consider the body is replaced by a blackbody i.e. $E \rightarrow E_b$ and $\alpha = 1$, the equation (1) becomes

$$E_b = q$$

Dividing eq. (1) by eq.(12),

$$\frac{E}{E_b} = \alpha \qquad (2)$$

At this point we may define emissivity, which is a measure of how good the body is an emitter as compared to blackbody. Thus the emissivity can be written as the ratio of the emissive power to that of a blackbody,

$$\frac{E}{E_b} = \epsilon \qquad (3)$$

Ratio of emissive power to the coefficient of absorption of any given wavelength is the same for all bodies at a given temperature and is equal to the emissive power of the black body at that temperature.

$$\frac{e_\lambda}{a_\lambda} = E \qquad (4)$$

4.25 STEFAN-BOLTZMANN LAW

According to this law the radiant energy (E) of the body is directly proportional to the fourth power of the temperature (T) of the body.

$$E \propto T^4$$

$$E = \sigma T^4$$

Where σ = Stefen constant, given by

$$\sigma = \frac{2\pi^5 {K_B}^4}{15h^3c^2}$$

4.26 WIEN'S DISPLACEMENT LAW

This law states that the product of then wavelength λ maximum energy and the absolute temperature (T) is a constant.

λ *mT Cons* tan *t*. This law shows that, as the temperature increases, the wavelength corresponding to maximum energy decreases.

Wien also showed that the maximum energy (Emax) is directly proportional to the fifth power of the absolute temperature. *E*m ax Dir Pro *T* 5, *E*m ax = *Cons* tan *tT* 5

By deducing this law he obtained a law called Wien's law of distribution of energy (Eλ), given by

$$E_\lambda = C_1\lambda^{-5}e^{-C_2/\lambda T}$$

Where C_1 and C_2 are constants given by

$$C_1 = 8\pi hc$$

$$C_2 = \frac{hc}{K_B}$$

4.27 NEWTON'S LAW OF COOLING

Newton's Law of Cooling states that the rate of change of the temperature of an object is proportional to the difference between its own temperature and the ambient temperature (i.e. the temperature of its surroundings).

When we apply the definition of Newton's Law of Cooling to an equation, we can get a formula. So, as per the law, the rate of a body cooling is in directly proportion to the difference in body's temperature.

We take body temperature as T and the surrounding temperature as T_0

The difference in temperature stays constant at 30^0 C. calculating the thermal energy Q. Note: the unit for thermal energy is joules. The formula derived from this equation is as follows:

$$\frac{dQ}{dT} \propto (T - T_0)$$

Where,

Mass of the body is represented by m, Specific heat is represented by s

And temperature is T, Surrounding temperature is T_0

The formula for thermal energy will be as follows:

$$Q = msT$$

Now let us calculate the rate of cooling.

$$\frac{dQ}{dt} = ms\frac{dT}{dt}$$

Therefore, we get,

$$ms\frac{dT}{dt} \propto (T - T_0)$$

Because we take mass and body heat as being constant, we can write the rate of change in temperature in the following manner:

$$\frac{dT}{dt} \propto (T - T_0)$$

The equation gives above demonstrates that with the increase in time, the difference in the temperature of the body and its surroundings, decreases, which means, the rate of temperature will also decrease.

Experimental verification of Newton's law of cooling

Newton’s law of cooling states that the rate of cooling of a body is directly proportional to the temperature difference between the body and the surroundings.The law holds good only for a small difference of temperature. Loss of heat by radiation depends on the nature of the surface and the area of the exposed surface.Let us consider a spherical calorimeter of mass m whose outer surface is blackened. It is filled with hot water of mass m_1. The calorimeter with a thermometer is suspended from a stand as shown in Figure.

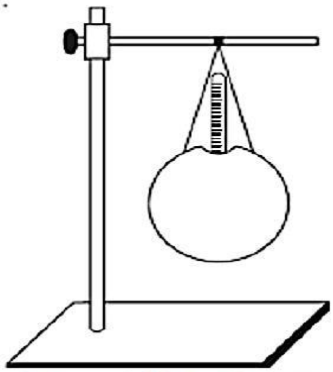


Fig. Newton's law of cooling

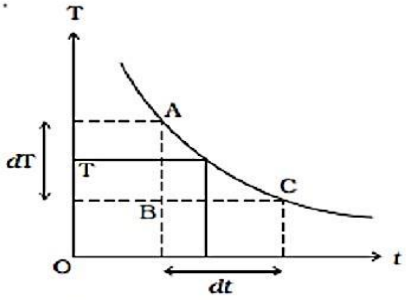


Fig. Cooling curve

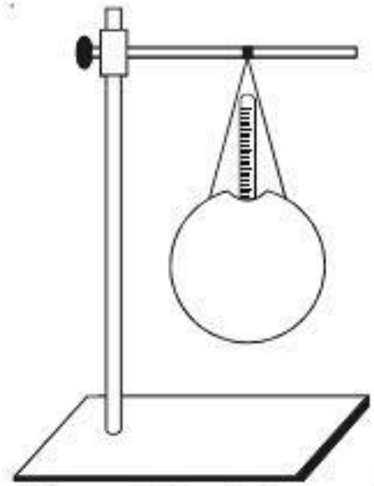


Fig. Newton's law of cooling

The calorimeter and the hot water radiate heat energy to the surroundings. Using a stop clock, the temperature is noted for every 30 seconds interval of time till the temperature falls by about 20° C. The readings are entered in a tabular column.

If the temperature of the calorimeter and the water falls from T^1 to T^2 in t seconds, the quantity of heat energy lost by radiation

$Q = (ms + m_1s_1)(T_1 / T_2)$, where s is the specific heat capacity of the material of the calorimeter and s1 is the specific heat capacity of water.

Rate of cooling = Heat energy lost / time taken

$Q / t = (ms + m_1s_1)(T_1 / T_2) / t$

If the room temperature is T_o, the average excess temperature of the calorimeter over that of the surroundings is $([(T_1 + T_2) / 2] - T_0)$

According to Newton's Law of cooling, $Q/t \propto ([(T_1 + T_2) / 2] - T_0)$

$(ms + m_1s_1)(T_1 ? T_2) / t[((T_1+T_2)/2) / T_0] = \text{constant}$

The time for every 4^o fall in temperature is noted. The last column in the tabular column is found to be the same. This proves Newtons Law of cooling.

1. Temperature range
2. Time t for every 4^o fall of temperature
3. Average excess of temperature $([(T_1 + T_2) / 2] - T_0)$
4. $([(T_1 + T_2) / 2] - T_0)t$

Temperature range	Time t for every 4^ofall of temperature	Average excess of temperature $\left(\frac{T_1+T_2}{2}-T_0\right)$	$\left(\frac{T_1+T_2}{2}-T_0\right)t$

A cooling curve is drawn by taking time along X-axis and temperature along Y-axis (Fig.).

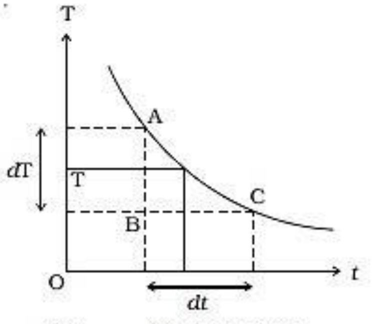


Fig. Cooling curve

From the cooling curve, the rate of fall of temperature at T is dT/dt = AB/BC

The rate of cooling dT/dt is found to be directly proportional to $(T - T_o)$. Hence Newton?s law of cooling is verified.

Just to know

Sadi Carnot (1796–1832): the "father" of Thermodynamics.

Nicolas Léonard Sadi Carnot (French: June 1796 – 24 August 1832) was a French military scientist and physicist, often described as the "father of thermodynamics". He published only one book, the Reflections on the Motive Power of Fire(Paris, 1824), in which he expressed, at the age of 27 years, the first successful theory of the maximum efficiency of heat engines.

Thermal Physics

PART – A

1. What are the entities responsible for thermal conduction of a solid?
2. Define coefficient of thermal conductivity.
3. Define thermal conductivity of heat.
4. What is meant by thermal resistance?
5. Mention four factors to be consider for providing good thermal effect for buildings.
6. What is meant by temperature gradient?
7. What are heat conduction and electrical conduction analogous to each other?
8. What are the characteristic of good and bad conductors?
9. Explain why the specimen used to determine thermal conductivity of a bad conductor should have a larger area and smaller thickness.
10. What are the three modes of heat transfer?
11. Define conduction of heat.
12. Define convection of heat.
13. Define radiation of heat.
14. Mention the method to determine thermal conductivity of good and bad conductor.
15. What is meant by Specific heat capacity?
16. What is meant by rate of rise / loss of heat?
17. What is an expansion joint?
18. What is a bimetallic strip?
19. What are heat exchangers?
20. What is the basic principle employed in Lee’s disc method for bad conductors?
21. Define refrigerator.

Thermal Physics

22. Define thermal diffusivity.
23. Define oven.
24. What is solar power?
25. Define Newton's Law of Cooling
26. State Wien’s displacement law.
27. Define co-efficient of Thermal conductivity and derive its units.
28. State the principle involved in Forbe’s method.
29. Explain why the specimen used to determine thermal conductivity of bad conductor should have larger area and smaller thickness.
30. In what direction does the growth of ice takes place?
31. Animal change their body to coiled or spherical shape during very cold weather.Why?
32. Why do we use glass in green house?
33. Among ice, water and steam, which has the highest thermal conductivity? Why?
34. Why metals are found to have high thermal conductivity?
35. Is there any heat conduction possible through vacuum?

PART –B

1. Describe the Forbe’s method to determine thermal conductivity of Metal with relevant theory.
2. Describe Lee’s disc method to find the co-efficient of thermal Conductivity of a bad conductor.
3. Derive the expression for heat conduction through a compoundMedia by series and parallel.
4. What is various classification of heat exchanger? Describe with Examples.

Thermal Physics

5. Explain briefly about the thermal insulation. Also discusses the importance of thermal insulation.
6. Describe the principle construction and working of microwave oven. Give its advantages and disadvantages.
7. Write notes on refrigeration. What are the various parts of therefrigerator and explain their function.
8. What is meant by solar power? Explain the working principle of solar water heater.
9. What you mean by Sea Breezes and Land Breezes.
10. Describe Prevost Theory of heat Exchange.
11. Differentiate between Heat absorption and Heat emission.
12. Write a note on Emissivity.
13. What is Black body radiation? What are the application of black body radiation?
14. State and explain the Stefan-Boltzmann law.
15. State and prove Newton’s law of cooling**.**
16. Explain how ice layer is formed on the surface of lakes, ponds, etc.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF PHYSICS

UNIT – V - PHYSICS FOR ENGINEERS-SPHA1101

UNIT 5

SENSORS AND DEVICES

.

5.1 INTRODUCTION

History has shown that advancements in materials science and engineering have been important drivers in the development of sensor technologies. For instance, the temperature sensitivity of electrical resistance in a variety of materials was noted in the early 1800s and was applied by **Wilhelm von Siemens in 1860** to develop a temperature sensor based on a copper resistor. The high resonance stability of single-crystal quartz, as well as its piezoelectric properties, have made possible an extraordinarily wide range of high performance, affordable sensors that have played an important role in everyday life and national defense. More recently, a new era in sensor technology was ushered in by the development of large-scale silicon processing, permitting the exploitation of silicon to create new methods for transducing physical phenomena into electrical output that can be readily processed by a computer.

Ongoing developments in materials technology will permit better control of material properties and behavior, thereby offering possibilities for new sensors with advanced features, such as greater fidelity, lower cost, and increased reliability. As noted in the preface, the Committee on New Sensor Technologies: Materials and Applications was asked to identify novel sensor materials that could benefit the manufacture and operation of advanced systems for the Department of Defense and the National Aeronautics and Space Administration and to identify research and development (R&D) efforts that could accelerate the development and incorporation of these emerging sensor materials in particular applications with potentially high payoff.

5.1.1 Definitions

Sensors and Devices

A sensor is a device that detects and responds to some type of input from the physical environment. The specific input could be light, heat, motion, moisture, pressure, or any one of a great number of other environmental phenomena.

A sensor is a device that receives a signal or stimulus and response with an electrical signal (Figure 5.1).

When input is a physical quantity and output electrical → Sensor

When input is electrical and output a physical quantity → Actuator

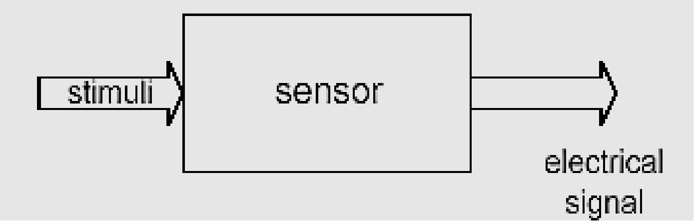


Fig 5.1 Sensor Device

Detectable Phenomenon

Stimulus	Quantity
Acoustic	Wave (amplitude, phase, polarization), Spectrum, Wave Velocity
Biological & Chemical	Fluid Concentrations (Gas or Liquid)
Electric	Charge, Voltage, Current, Electric Field (amplitude, phase, polarization), Conductivity, Permittivity
Magnetic	Magnetic Field (amplitude, phase, polarization), Flux, Permeability
Optical	Refractive Index, Reflectivity, Absorption
Thermal	Temperature, Flux, Specific Heat, Thermal Conductivity
Mechanical	Position, Velocity, Acceleration, Force, Strain, Stress, Pressure, Torque

5.1.1 Need For Sensors

Sensors and Devices

Sensors are omnipresent. They embedded in our bodies, automobiles, airplanes, cellular telephones, radios, chemical plants, industrial plants and countless other applications. Without the use of sensors, there would be no automation!! Imagine having to manually fill Poland Spring bottles.

5..1.2 A good sensor obeys the following rules:

- it is sensitive to the measured property,
- it is insensitive to any other property likely to be encountered in its application,
- it does not influence the measured property.

5.1.3 Types Of Sensors

- **Direct:** A sensor that can convert a non-electrical stimulus into an electrical signal with intermediate stages, e.g. Thermocouple (temperature to voltage)

- **Indirect:** A sensor that multiple conversion steps to transform the measured signal into an electrical signal, for example a fiber-optic displacement sensor (Light Current →photons →current)

5.2 CLASSIFICATION OF SENSORS

Based on physical laws or convenient distinguishing property

1. Active and Passive sensors
2. Contact and non-contact sensors
3. Absolute and relative sensors
4. Others

5.2.1 Active and Passive Sensors

Active sensor: A sensor that requires external power to operate, e.g. carbon microphone, thermistors, strain gauges, capacitive and inductive sensors, etc. The active sensor is also called as parametric sensor (output is a function of a parameter-like resistance).

Passive sensor: It generates its own electric signal and does not require a power source, e.g. thermocouples, magnetic microphones, piezo electric sensors, photodiode. Also called as self-generating sensors

5.2.2 Contact and Non-Contact Sensors

Contact sensor:A sensor that requires physical contact with the stimulus, e.g. strain gauges, temperature sensors.

Non-contact sensor: It requires no physical contact, e.g. most optical and magnetic sensors, infrared thermometers, etc.

5.2.3 Absolute and Relative Sensors

Absolute sensor:A sensor that reacts to a stimulus on an absolute scale, such as thermistors, strain gauges, etc., (thermistor always reads the absolute temperature).

Relative scale:The stimulus is sensed relative to a fixed or variable reference, for example thermocouple measures the temperature difference; pressure is often measured relative to atmospheric pressure.

5.2.4 Others

Classification based on broad area of detection: Electric sensors, Magnetic, Electromagnetic, Acoustic, Chemical, Optical, Heat, Temperature, Mechanical, Radiation, Biological etc.

Classification based on physical law: Photoelectric, Magneto electric, Thermoelectric, Photoconductive, Photo magnetic, Thermomagnetic, Thermo optic, Electro chemical, Magneto resistive, Photo elastic etc.

Classification based on Specification: Accuracy, sensitivity, Stability, response time, hysteresis, Frequency response, input, resolution, linearity, hardness, cost, size, weight, conduction materials, temperature etc.

***Classification as per Application**:* Agriculture, Automotive, Civil engineering and construction, Domestic appliances, Commerce, Finance Environment, Meteorology, security, Energy, Information and Telecommunication, Health and medicine, Marine, Military and Space, Recreation and toys, Scientific measurement, Manufacturing and Transportation and many more…

5.3 PRINCIPLE OF OPERATION

5.2.1 Linear and Rotational Sensors

Linear and rotational position sensors are two of the most fundamental of all measurements used in a typical mechatronics system. In general, the position sensors produce an electrical output that is proportional to the displacement they experience. There are contact type sensors such as strain gage, tachometer, etc. The noncontact type includes encoders, hall effect, capacitance, inductance, and interferometer type.

5.2.2 Acceleration Sensors

Measurement of acceleration is important for systems subject to shock and vibration. Although acceleration can be derived from the time history data obtainable from linear or rotary sensors, the accelerometers whose output is directly proportional to the acceleration is preferred. Two common types include the *seismic mass* type and the *piezoelectric* accelerometer. The seismic mass type accelerometer is based on the relative motion between a mass and the supporting structure. The natural frequency of the seismic mass limits its use too low to medium frequency applications. The piezoelectric accelerometer, however, is compact and more suitable for high frequency applications.

5.2.3 Force, Torque, and Pressure Sensors

Sensors and Devices

Among many type of force/torque sensors, the *strain gage dyanamometers* and *piezoelectric type* are most common. Both are available to measure force and/or torque either in one axis or multiple axes. The dynamometers make use of mechanical members that experiences elastic deflection when loaded. These types of sensors are limited by their natural frequency. On the other hand, the piezoelectric sensors are particularly suitable for dynamic loadings in a wide range of frequencies. They provide high stiffness, high resolution over a wide measurement range, and are compact.

5.2.4 Flow Sensors

Flow sensing is relatively a difficult task. The fluid medium can be liquid, gas, or a mixture of the two. Furthermore, the flow could be laminar or turbulent and can be a time-varying phenomenon. The *venture meter* and *orifice plate* restrict the flow and use the pressure difference to determine the flow rate. The *pitot tube* pressure probe is another popular method of measuring flow rate. When positioned against the flow, they measure the total and static pressures. The flow velocity and in turn the flow rate can then be determined. The *rotameter* and the *turbine meters* when placed in the flow path, rotate at a speed proportional to the flow rate. The *electromagnetic flow meters* use noncontact method. Magnetic field is applied in the transverse direction of the flow and the fluid acts as the conductor to induce voltage proportional to the flow rate. *Ultrasonic flow meters* measure fluid velocity by passing high-frequency sound waves through fluid.

5.2.5 Temperature Sensors

A variety of devices are available to measure temperature, the most common of which are thermocouples, thermisters, resistance temperature detectors (RTD), and infrared types. *Thermocouples* are the most versatile, inexpensive, and have a wide range (up to 1200 ∞ C typical). A thermocouple simply consists of two dissimilar metal wires joined at the ends to create the sensing junction. When used in conjunction with a reference junction, the temperature difference between the reference junction and the actual temperature shows up as a voltage potential. *Thermisters* are semiconductor devices whose resistance changes as the temperature changes. They are good for very high sensitivity measurements in a limited range of up to 100 ∞ C. The relationship between the temperature and the resistance is nonlinear. The *RTD* s use the phenomenon that the resistance of a metal changes with temperature. They are, however, linear over a wide range and most stable. *Infrared type* sensors use the radiation heat to sense the temperature from a distance. These noncontact sensors can also be used to sense a field of vision to generate a thermal map of a surface.

5.2.6 Proximity Sensors

They are used to sense the proximity of an object relative to another object. They usually provide a on or off signal indicating the presence or absence of an object. *Inductance, capacitance*, *photoelectric*, and *hall effect* types are widely used as proximity sensors. Inductance proximity sensors consist of a coil wound around a soft iron core. The inductance of the sensor changes when a ferrous object is in its proximity. This change is converted to a voltage-triggered switch. Capacitance types are similar to inductance except the proximity of an object changes the gap and affects the capacitance. Photoelectric sensors are normally aligned with an infrared light source. The proximity of a moving object interrupts the light beam causing the voltage level to change. Hall effect voltage is produced when a current-carrying conductor is exposed to a transverse magnetic field. The voltage is proportional to transverse distance between the hall effect sensor and an object in its proximity.

5.2.7 Light Sensors

Light intensity and full field vision are two important measurements used in many control applications. *Phototransistors, photoresistors,* and *photodiodes* are some of the more common type of light intensitysensors. A common photoresistor is made of cadmium sulphide whose resistance is maximum when thesensor is in dark. When the photoresistor is exposed to light, its resistance drops in proportion to theintensity of light.

5.2.8 Smart Material Sensors

There are many new smart materials that are gaining more applications as sensors, especially in distributed sensing circumstances. Of these, *optic fibers, piezoelectric,* and *magnetostrictive* materials have found applications. Within these, optic fibers are most used. Optic fibers can be used to sense strain, liquid level, force, and temperature with very high resolution. Since they are economical for use as *in situ* distributed sensors on large areas, they have found numerous applications in smart structure applications such as damage sensors, vibration sensors, and cure-monitoring sensors. These sensors use the inherent material (glass and silica) property of optical fiber to sense the environment

5.2.9 Micro and Nanosensors

Microsensors (sometimes also called MEMS) are the miniaturized version of the conventional macrosensors with improved performance and reduced cost. Silicon micromachining technology has helped the development of many microsensors and continues to be one of the most active research and development topics in this area. Vision microsensors have found applications in medical technology. A *fiberscope* of approximately 0.2 mm in diameter has been developed to inspect flaws inside tubes. Another example is a *microtactile sensor* which uses laser light to detect the contact between a catheter and the inner wall of blood vessels during insertion that has sensitivity in the range of 1 mN. Similarly, the progress made in the area of nanotechnology has fuelled the development of nanosensors. These are relatively new sensors that take one step further in the direction of miniaturization and are expected to open new avenues for sensing applications.

5.3 PRESSURE SENSOR

5.3.1 Introduction

A **pressure sensor** measures pressure, typically of gases or liquids. Pressure is an expression of the force required to stop a fluid from expanding, and is usually stated in terms of force per unit area. A pressure sensor usually acts as a transducer; it generates a signal as a function of the pressure imposed. For the purposes of this article, such a signal is electrical. Pressure sensors are used for control and monitoring in thousands of everyday applications. Pressure sensors can also be used to indirectly measure other variables such as fluid/gas flow, speed, water level, and altitude. Pressure sensors can alternatively be called **pressure transducers**, **pressure transmitters**, **pressure senders**, **pressure indicators**, **piezometers** and **manometers**, among other names. Pressure sensors can vary drastically in technology, design, performance, application suitability and cost. A conservative estimate would be that there may be over 50 technologies and at least 300 companies making pressure sensors worldwide.

5.3.2 Types of Pressure Measurements

Pressure sensors can be classified in terms of pressure ranges they measure, temperature ranges of operation, and most importantly the type of pressure they measure. Pressure sensors are variously named according to their purpose, but the same technology may be used under different names.

- ***Absolute pressure sensor***

  This sensor measures the pressure relative to perfect vacuum.

- ***Gauge pressure sensor***

  This sensor measures the pressure relative to atmospheric pressure. A tire pressure gauge is an example of gauge pressure measurement; when it indicates zero, then the pressure it is measuring is the same as the ambient pressure.

- ***Vacuum pressure sensor***

  It may be used to describe a sensor that measures pressures below atmospheric pressure, showing the difference between that low pressure and atmospheric pressure (i.e. negative gauge pressure), but it may also be used to describe a sensor that measures low pressure relative to perfect vacuum (i.e. absolute pressure).

- ***Differential pressure sensor***

  This sensor measures the difference between two pressures, one connected to each side of the sensor. Differential pressure sensors are used to measure many properties, such as pressure drops across oil filters or air filters, fluid levels (by comparing the pressure above and below the liquid) or flow rates (by measuring the change in pressure across a restriction).

- ***Sealed pressure sensor***

  This sensor is similar to a gauge pressure sensor except that it measures pressure relative to some fixed pressure rather than the ambient atmospheric pressure (which varies according to the location and the weather).

5.3.3 There are Many Applications for Pressure Sensors

- ***Pressure sensing***

  This is where the measurement of interest is pressure, expressed as a force per unit area. This is useful in weather instrumentation, aircraft, automobiles, and any other machinery that has pressure functionality implemented.

- ***Altitude sensing***

  This is useful in aircraft, rockets, satellites, weather balloons, and many other applications. All these applications make use of the relationship between changes in pressure relative to the altitude.

Sensors and Devices

- ***Flow sensing***

  This is the use of pressure sensors in conjunction with the venturi effect to measure flow. Differential pressure is measured between two segments of a venturi tube that have a different aperture.

- ***Level / depth sensing***

  A pressure sensor may also be used to calculate the level of a fluid. This technique is commonly employed to measure the depth of a submerged body (such as a diver or submarine), or level of contents in a tank (such as in a water tower **Leak testing**

- ***Ratiometric Correction of Transducer Output***

  Piezoresistive transducers configured as Wheatstone bridges often exhibit ratio metric behavior with respect not only to the measured pressure, but also the transducer supply voltage.

5.4 BOURDON TUBE PRESSURE GAUGE

5.4.1 Introduction

- Bourdon Tubes are known for its very high range of differential **pressure measurement** in the range of almost 100,000 psi (700 MPa). It is an elastic type pressure transducer.
- The device was invented by Eugene Bourdon in the year 1849.
- The basic idea behind the device is that, cross-sectional tubing when deformed in any way will tend to regain its circular form under the action of pressure.
- The bourdon pressure gauges used today have a slight elliptical cross-section and the tube is generally bent into a C-shape or arc length of about 27 degrees.
- In figure, the pressure input is given to a socket which is soldered to the tube at the base. The other end or free end of the device is sealed by a tip. This tip is connected to a segmental lever through an adjustable length link.

- The lever length may also be adjustable. The segmental lever is suitably pivoted and the spindle holds the pointer as shown in the figure. A hair spring is sometimes used to fasten the spindle of the frame of the instrument to provide necessary tension for proper meshing of the gear teeth and thereby freeing the system from the backlash.
- Any error due to friction in the spindle bearings is known as lost motion. The mechanical construction has to be highly accurate in the case of a Bourdon Tube Gauge. If we consider a cross-section of the tube, its outer edge will have a larger surface than the inner portion. The tube walls will have a thickness between 0.01 and 0.05 inches.

5.4.2 Working

As the fluid pressure enters the bourdon tube, it tries to be reformed and because of a free tip available, this action causes the tip to travel in free space and the tube unwinds. The simultaneous actions of bending and tension due to the internal pressure make a non-linear movement of the free tip. This travel is suitable guided and amplified for the measurement of the internal pressure. But the main requirement of the device is that whenever the same pressure is applied, the movement of the tip should be the same and on withdrawal of the pressure the tip should return to the initial point.

A lot of compound stresses originate in the tube as soon as the pressure is applied. This makes the travel of the tip to be non-linear in nature. If the tip travel is considerably small, the stresses can be considered to produce a linear motion that is parallel to the axis of the link. The small linear tip movement is matched with a rotational pointer movement. This is known as multiplication, which can be adjusted by adjusting the length of the lever. For the same amount of tip travel, a shorter lever gives larger rotation. The approximately linear motion of the tip when converted to a circular motion with the link-lever and pinion attachment, a one-to-one correspondence between them may not occur and distortion results. This is known as angularity which can be minimized by adjusting the length of the link.

Other than C-type, bourdon gauges (Figure 5.2) can also be constructed in the form of a helix or a spiral. The types are varied for specific uses and space accommodations, for better linearity and larger sensitivity. For thorough repeatability, the bourdon tubes materials must have good elastic or spring characteristics. The surrounding in which the process is carried out is also important as corrosive atmosphere or fluid would require a material which is corrosion proof. The commonly used materials are phosphor-bronze, silicon-bronze, beryllium-copper, inconel, and other C-Cr-Ni-Mo alloys, and so on.

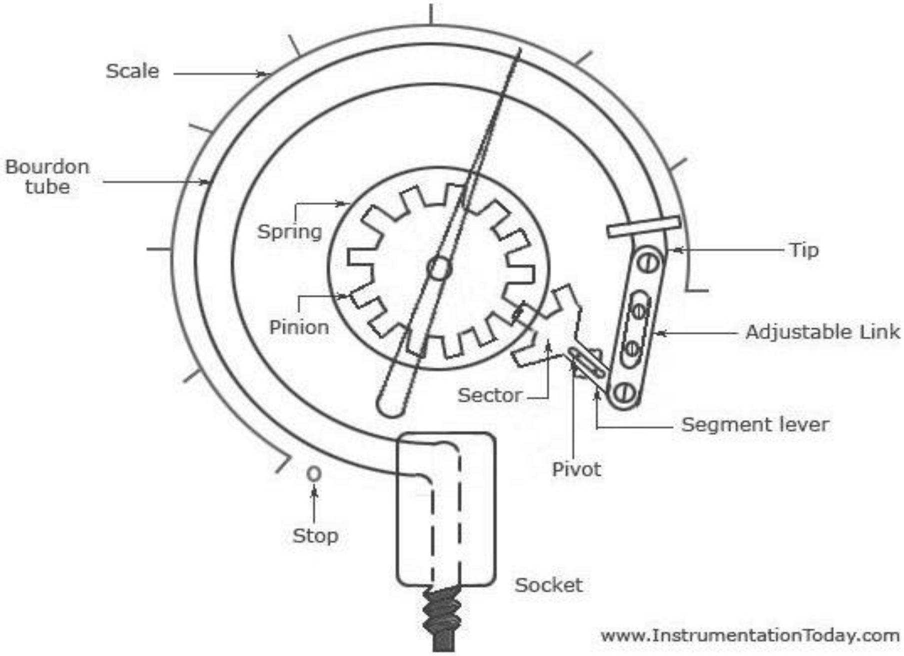
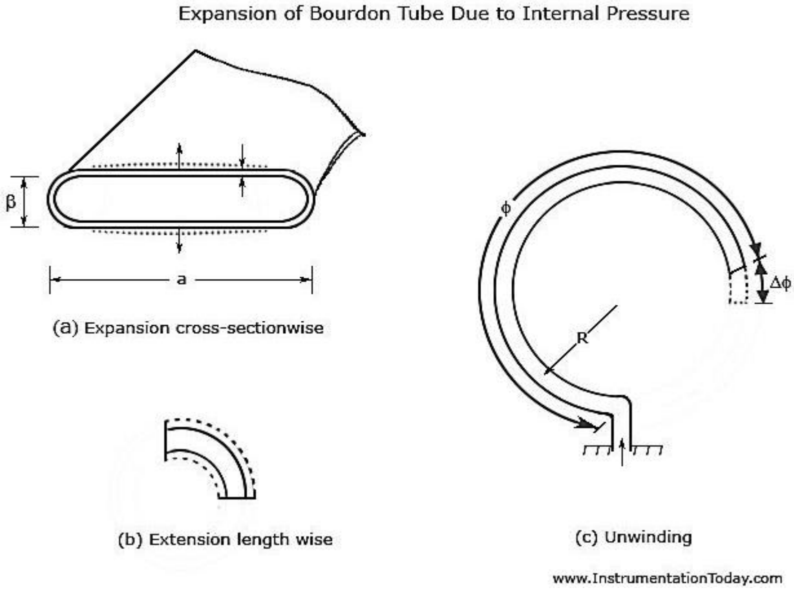


Fig 5.2 Bourdon Tube Pressure Gauge

In the case of forming processes, empirical relations are known to choose the tube size, shape and thickness and the radius of the C-tube. Because of the internal pressure, the near elliptic or rather the flattened section of the tube tries to expand as shown by the dotted line in the figure 5.3 below (a). The same expansion lengthwise is shown in figure (b). The arrangement of the tube, however forces an expansion on the outer surface and a compression on the inner surface, thus allowing the tube to unwind. This is shown in figure (c).

5.4.3 Expansion Of Bourdon Tube Due To Internal Pressure

Like all elastic elements a bourdon tube also has some hysteresis in a given pressure cycle. By proper choice of material and its heat treatment, this may be kept to within 0.1 and 0.5 percent of the maximum pressure cycle. Sensitivity of the tip movement of a bourdon element without restraint can be as high as 0.01 percent of full range pressure reducing to 0.1 percent with restraint at the central pivot.



(a) Expansion cross-sectionwise

(b) Extension length wise

(c) Unwinding

www.InstrumentationToday.com

Fig.5.3 Expansion of Bourdon Tube due to Internal Pressure

5.5 TEMPERATURE SENSOR /THERMAL SENSOR/ THERMODYNAMIC SENSOR

Temperature is the most often measured environmental quantity. The physical, chemical, mechanical and biological systems are affected by temperature. Certain chemical reaction, biological process and even electronic circuits perform best within limited temperature ranges. Temperature sensing can be done either through direct contact with the heating source or without direct contact with the source using radiated energy instead.

Types

There is wide variety of temperature sensors, including Thermocouples, Resistance temperature detector (RTD), Thermistor, Infrared & Semiconductor Sensors.

Temperature sensor classification

(i) *Primary Sensor:* If it is possible to relate the temperature (T) directly in the form of Q=NU (Q> any physical quantity, N-number and in unit (U)).eg: Gas sensor, Pressure sensor, Vapour sensor etc.

(ii) Secondary Sensor: If the relation between the temperature (T) with Q,N and U is not direct. eg: Thermal expansion types- solid, liquid and gas; Resistance thermometer- thermocouple, thermistor etc.

Application

Monitoring: Portable Equipment
CPU Temperature
Battery Temperature
Ambient Temperature

Compensation: Oscillator Drift in Cellular Phones
Thermocouple cold-junction compensation

Control: Battery Charging
Process control

5.5.1 Resistance Type Thermal Sensor

5.5.1.1 *Thermocouple*

Thermoelectric sensor works on the principle of Seebeck effect and Which is made by joining of two dissimilar metal wires (A,B) are welded together at its ends thus forming 2 junctions (J1 & J2). The two junctions are maintained at two different temperatures (T1 & T2) then an emf (electro motive force) is generated between the 2 junctions because of temperature difference. This is called as Thermo emf/ Tic potential. This effect is called Seebeck effect. The emf appears across the free ends of the thermocouple wires, where it is measured and converted into units of heat calibration. Thermocouples are suitable to be used in temperature ranges from -450F to -4200 F. The magnitude of thermo emf depends on the, (i) composition of two metal wires (ii) temperature T1 & T2 of the respective junctions J1 & J2 (Figure 5.4).

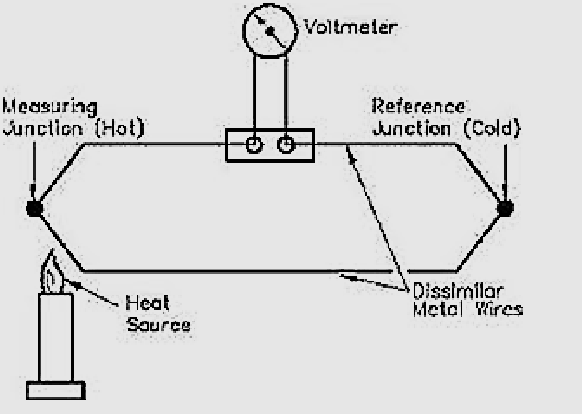


Fig.5.4 Thermocouple

Peltier effect: It states that in a thermocouple, if a current (I) is allowed to flow in the circuit, heat is generated at the cold junction and is absorbed at the hot junction. This effect is known as Peltier effect. Heat flow is proportional to the current I in the circuit.

$$\mathbf{Hf = \pi I}$$

π – Peltier co-efficient.

Thomson effect: With a current flowing through a single metal wire, its heat content changes and a temperature gradient exists along the length. Heat flow is proportional to the current I and temperature gradient.

$$\mathbf{Hf = \sigma\, I\, .\Delta T}$$

σ – Peltier co-efficient

ΔT- temperature gradient.

Types of Thermocouples

1) Base metal type: consists of elements metal & their alloys they are identified by letters E,J, K, N, T. Examples: Chromel Alumel, Chromel constantan, Copper constantan.
2) Nobble/ Precious metal type: made from nobble metal & their alloys. They are identified by letters G, C, D, B, R, S. Example: pt(6%)-Rhodium, pt(30%)-Rhodium.
3) Non-metallic type.

Important Factors

High thermoelectric power, low electrical resister, linearity of E-T curve over the range, high melting point, material should be available as pure & homogeneous, should be stable.

Applications

- Thermocouples are suitable for use in industrial furnaces Applications
- Plastic injection molding machinery
- Food processing equipment
- Deicing
- Semiconductor processing
- Heat treating
- Medical equipment
- Industrial heat treating
- Packaging equipment

Advantages

- Simple, Rugged
- High temperature operation
- Low cost
- No resistance lead wire problems
- Point temperature sensing
- Fastest response to temperature changes

Disadvantages

- Least stable, least repeatable
- Low sensitivity to small temperature changes
- Extension wire must be of the same thermocouple type
- Wire may pick up radiated electrical noise if not shielded
- Lowest accuracy

Limitation

- Non linear
- Low voltage
- Reference required
- Least stable
- Least sensitive output vs. temperature change.

5.5.2.2 Thermistor

Introduction

The thermistors are made up of ceramic like semiconducting materials. They are mostly composed of oxides of manganese, nickel and cobalt having the resistivities if about 100 to 450,000 ohm-cm. Since the resistivity of the thermistors is very high the resistance of the circuit in which they are connected for measurement of temperature can be measured easily. This resistance is calibrated against, the input quantity, which is the temperature, and its value can be obtained easily. Thermistors are available in various shapes like disc, rod, washer, bead etc. They are of small size and they all can be fitted easily to the body whose temperature has to be measured and also can be connected to the circuit easily. Most of the thermistors are quite cheap.

Principle of Working of Thermistors

As mentioned earlier the resistance of the thermistors decreases with the increase its temperature (Fig. 5.5). The resistance of thermistor is given by:

$$R = R_o e^k$$

$$K = \beta(1/T - 1/T_o)$$

Where R is the resistance of the thermistor at any temperature T in ^oK (degree Kelvin)

R_o is the resistance of the thermistors at particular reference temperature T_o in ^oK

e is the base of the Naperian logarithms

β is a constant whose value ranges from 3400 to 3900 depending on the material used for the thermistors and its composition.

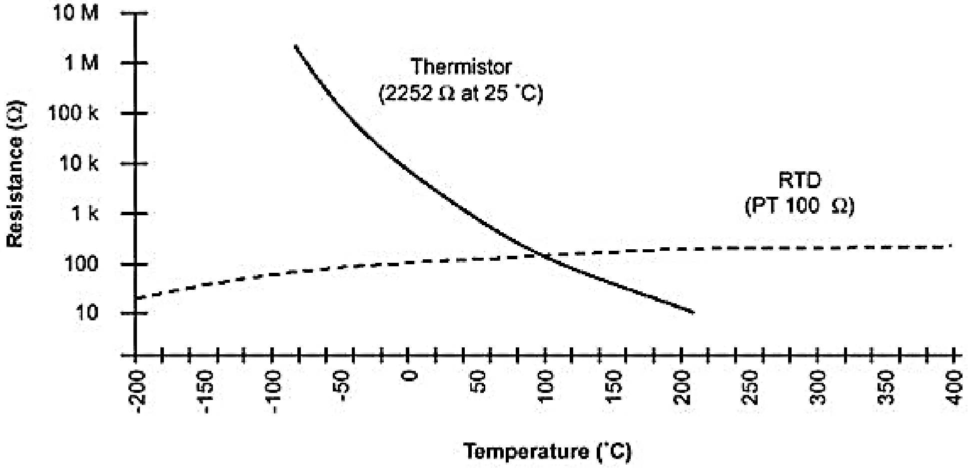


Fig 5.5 Thermistor operation

The thermistor acts as the temperature sensor and it is placed on the body whose temperature is to be measured. It is also connected in the electric circuit. When the temperature of the body changes, the resistance of the thermistor also changes, which is indicated by the circuit directly as the temperature since resistance is calibrated against the temperature. The thermistor can also be used for some control which is dependent on the temperature.

Types of Thermistors

Thermistors are classified into two types such as

- Thermistor Elements
- Thermistor Probes

Thermistor Elements

Thermistor elements are the simplest form of thermistor, it is commonly used when space is very limited. OMEGA offers a wide variety of thermistor elements which vary not only in form factor, but also in their resistance Vs temperature characteristics. Since thermistors are non-linear, the device used to read the temperature must linearize the reading.

Thermistor Probes

The standalone thermistor element is comparatively delicate and cannot be located in a rugged environment. OMEGA offers thermistor probes that are thermistor elements fixed in metal tubes. Thermistor probes are much more suitable for industrial environments than thermistor elements.

Advantages of Thermistors

1) When the resistors are connected in the electrical circuit, heat is dissipated in the circuit due to flow of current. This heat tends to increase the temperature of the resistor due to which their resistance changes. For the thermistor the definite value of the resistance is reached at the given ambient conditions due to which the effect of this heat is reduced.

2) In certain cases even the ambient conditions keep on changing, this is compensated by the negative temperature characteristics of the thermistor. This is quite convenient against the materials that have positive resistance characteristics for the temperature.

3) The thermistors are used not only for the measurement of temperature, but also for the measurement of pressure, liquid level, power etc.

4) They are also used as the controls, overload protectors, giving warnings etc.

5) The size of the thermistors is very small and they are very low in cost. However, since their size is small they have to be operated at lower current levels.

Applications of thermistors

- A Thermistor is used to measure the temperature.
- The thermistor is used as an electrical circuit component
- For temperature compensation
- Circuit protection
- Voltage regulation
- Time delay, and Volume control.
- Thermistors are used in an automotive applications
- Instrumentation and Communication
- Consumer electronics
- Food handling and processing
- Industrial electronics
- Medical electronics
- Military and aerospace

5.6 VIBRATION SENSORS

Vibration is one of the most popular phenomena that exists in our daily life, which is everywhere and at all the time. Vibration is generated as a result of mechanical disturbance from sources such as music/sound, noise, engine, wind and many more. Detection of vibration is an important sensor technology for monitoring the operation of machines, bridges and buildings, warrant of security, prediction of natural disasters and more. As we know, the vibration sensor testing technology has been developed gradually from early last century. With scientists' exploring and researching, and accordingly the test methods and the types of sensors are evolving and maturing. Vibration measurements usually include vibration displacement, velocity, acceleration and others' measurement, usually, the device that converse the vibration into the electrical is called as vibration sensor. Especially in recent years, vibration measurement has become an important method in mechanical structural product's research, design, produce, apply and maintenance. Thus, a variety of vibration sensors made by the effect of physical have drawn more and more attention, with the development of computer technology, electronic technology and manufacturing process, a variety of vibration sensors have come forth in succession in order to using in different areas (Fig. 5.7).

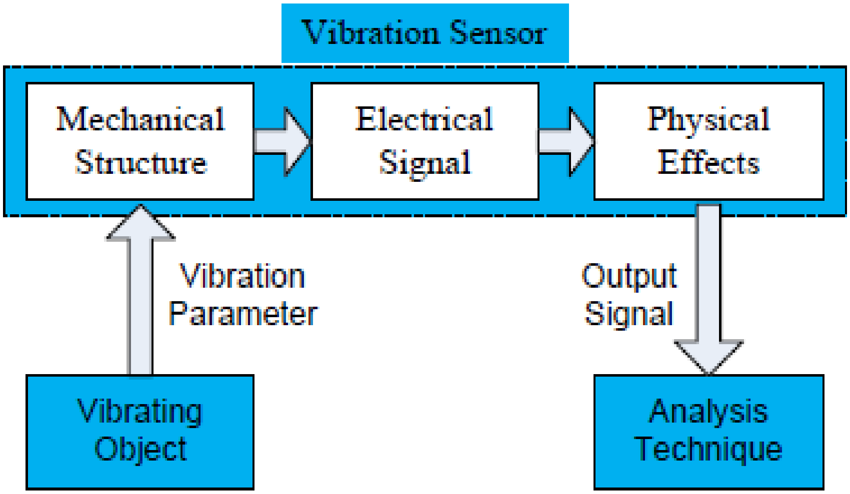


Fig.5.6 Basic Measurement Principle of Vibration Sensor

Sensors and Devices

1) Measurements on Structures or Machinery Casings: Accelerometers and Velocity Sensors

Used in gas turbines, axial compressors, small and mid-size pumps. These sensors detect high frequency vibration signals related to bearing supports, casing and foundation resonances, vibration in turbine/compressor vanes, defective roller or ball bearings, noise in gears, etc.

2) Displacement measurements relative to rotating shafts: Proximity Probes (capacitance or eddy-current)

Used in turbomachinery supported on fluid film bearings, centrifugal compressors, gears and transmissions, electric motors, large pumps (>300HP), some turbines and fans. These sensors detect shaft static displacements, unbalance response, misalignment, shaft bending, excessive loads in bearings, dynamic instabilities, etc.

(i)Velocity Sensors

Electromagnetic linear velocity transducers: Typically used to measure oscillatory velocity. A permanent magnet moving back and forth within a coil winding induces an *emf* in the winding.

This *emf* is proportional to the velocity of oscillation of the magnet. This permanent magnet maybe attached to the vibrating object to measure its velocity.

Electromagnetic tachometer generators: Used to measure the angular velocity of vibrating objects. They provide an output voltage/frequency that is proportional to the angular velocity. *DC tachometers* use a permanent magnet or magneto, while the *AC tachometers* operate as a variable coupling transformer, with the coupling coefficient proportional to the rotary speed.

Advantages	***Disadvantages***
Simple to install	Low resonant frequency & phase shift
Good response	Cross noise
Stand high temperature	Big and heavy
Do not require external power	Require electronic integration for displacement
Lowest cost	

(ii)Acceleration Sensors

Capacitive accelerometers: Used generally in those that have diaphragm supported seismic mass as a moving electrode and one/two fixed electrodes. The signal generated due to change incapacitance is post-processed using LC circuits etc., to output a measurable entity.

Piezoelectric accelerometers: Acceleration acting on a seismic mass exerts a force on the piezoelectric crystals, which then produce a proportional electric charge. The piezoelectric crystals are usually preloaded so that either an increase or decrease in acceleration causes a change in the charge produced by them. But they are not reliable at very low frequencies.

Potentiometric accelerometers: Relatively cheap and used where slowly varying acceleration is to be measured with a fair amount of accuracy. In these, the displacement of a spring mass system is mechanically linked to a viper arm, which moves along a potentiometric resistive element. Various designs may have either viscous, magnetic or gas damping.

Reluctive accelerometers: They compose accelerometers of the differential transformer type or the inductance bridge type. The AC outputs of these vary in phase as well as amplitude. They are converted into DC by means of a phase-sensitive demodulator.

Servo accelerometers: These use the closed loop servo systems of force-balance, torque-balance or null-balance to provide close accuracy. Acceleration causes a seismic mass to move. The motion is detected by one of the motion-detection devices, which generate a signal that acts as an error signal in the servo-loop. The demodulated and amplified signal is then passed through a passive damping network and then applied to the torquing coil located at the axis of rotation of the mass. The torque is proportional to the coil current, which is in turn proportional to the acceleration.

Strain Gage accelerators: these can be made very small in size and mass. The displacement of the spring-mass system is converted into a change in resistance, due to strain, in four arms of a Wheatstone bridge. The signal is then post-processed to read the acceleration.

Advantages	***Disadvantages***
Simple to install	Sensitive to high frequency noise
Good response at high frequencies	Require external power
Stand high Temperature	Require electronic integration for velocity displacement
Small size and	

5.6.1 Optic Fiber Sensor

Generally by the optical fiber sensor, laser and light detector composed of the three parts of Optic fiber sensors. According to the different Operating Principle of Optical fiber sensor can be divided into functional and non functional. The former is the use of the characteristics of the fiber itself, and use the optical fiber as the sensitive components. The latter is the use of other sensitive components to detect changes of the measured physical quantity; just optical fiber is used as transmission medium to transport the optical signal from distant or inaccessible location of. In practice, the optical fiber as the sensitive components of vibration information directly is difficult to separate the impact of changes from other physical quantities, therefore, the non-functional optical fiber vibration sensors is widely used in the field of vibration detection, where the basic principle is the use of other sensitive detect changes of the measured physical quantity, and the light parameter is modulated by sensitive components. Figure 5.7 is a phase-modulated optical fiber vibration sensor schematic diagram. The vibrating object change the relative phase of signal beam and reference beam, which result in a phase modulation, and by demodulation and detection the phase modulation, you can get the corresponding vibration amplitude.

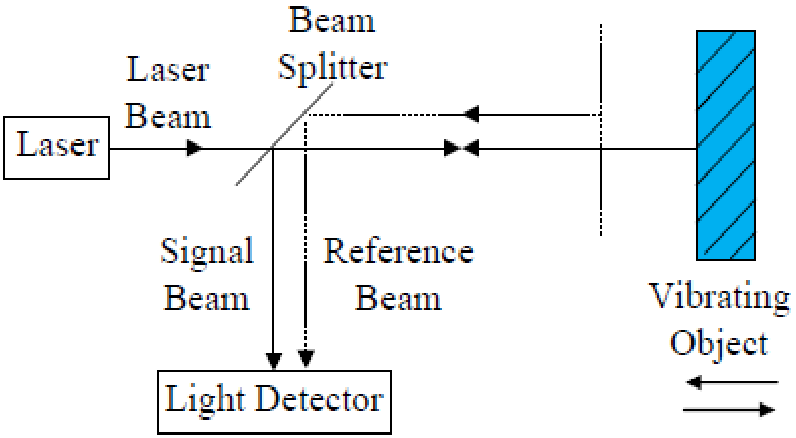


Fig.5.7 Schematic diagram of a phase-modulated optical fiber

Vibration Sensor

The prominent features of the optical fiber sensor conclude their ability to be lightweight, very small size, high sensitivity, fast response, resistance to electromagnetic interference, corrosion resistance, electrical insulation, soft bend, suitable for long-distance transmission, and easy to connect with the computer and make telemetry network with fiber optic transmission systems, especially can long-distance vibration in harsh industrial environments. Practice has proved that it has high sensitivity and reliability of persistent work, which can detect the vibration amplitude from 10-12 meters and can be used to three-dimensional vibration measurements.

The disadvantages conclude the narrow range of measurement frequency, high cost and unfamiliarity to the end user. Therefore, the optical fiber vibration sensor has the broad value of further research and development.

5.6.2Geophone

A **geophone** is a device that converts ground movement (velocity) into voltage, which may be recorded at a recording station. The deviation of this measured voltage from the base line is called the seismic response and is analyzed for structure of the earth. The majority of geophones are used in reflection seismology to record the energy waves reflected by the subsurface geology. In this case the primary interest is in the vertical motion of the Earth's surface. However, not all the waves are upwards travelling. A strong, horizontally transmitted wave known as ground-roll also generates vertical motion that can obliterate the weaker vertical signals. By using large areal arrays tuned to the wavelength of the ground-roll the dominant noise signals can be attenuated and the weaker data signals reinforced (Figure 5.8).

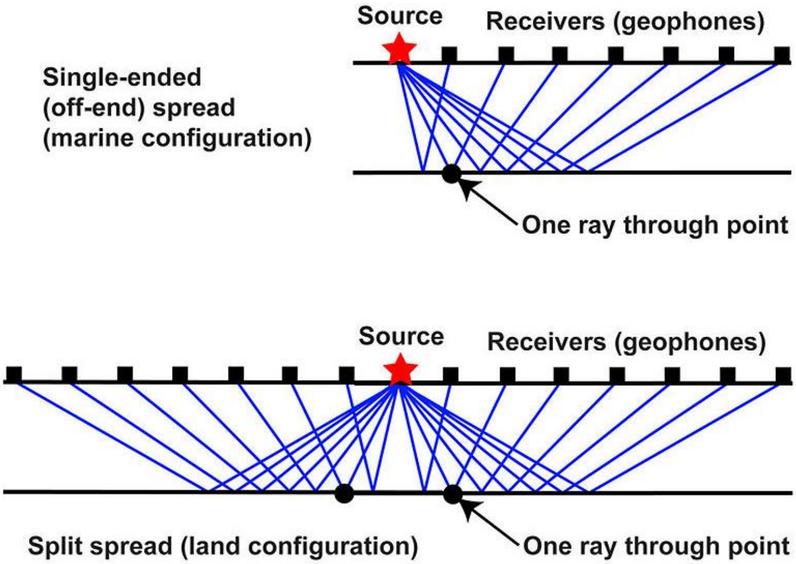


Fig 5.8 Geophone operation

5.7 ACOUSTIC SENSORS

Acoustic sensors detect sound via means of using microphones, or other types of filters. Surface acoustic wave sensors are a class of microelectromechanical systems (MEMS) which rely on the modulation of surface acoustic waves to sense.The most common sensor used for acoustic measurement is the microphone. Measurement-grade microphones are different from typical recording-studio microphones because they can provide a detailed calibration for their response and sensitivity. Other sensors include hydrophones for measuring sound in water or accelerometers for measuring vibrations causing sound.

Microphones

- Microphone: acoustic sensors for air waves in the audible range
- Hydrophone: acoustic sensor for liquid wavesmicrophone / hydrophone are pressure sensors with a wide dynamic range.
- A microphone / hydrophone is a pressure transducer, adapted for the transduction of sound / liquid waves.
- All microphones / hydrophones have a moving diaphragm and a displacement transducer that converts this motion into an electric signal.
- Microphones / hydrophones differ by: sensitivity, direction characteristics, frequency bandwidth, dynamic range

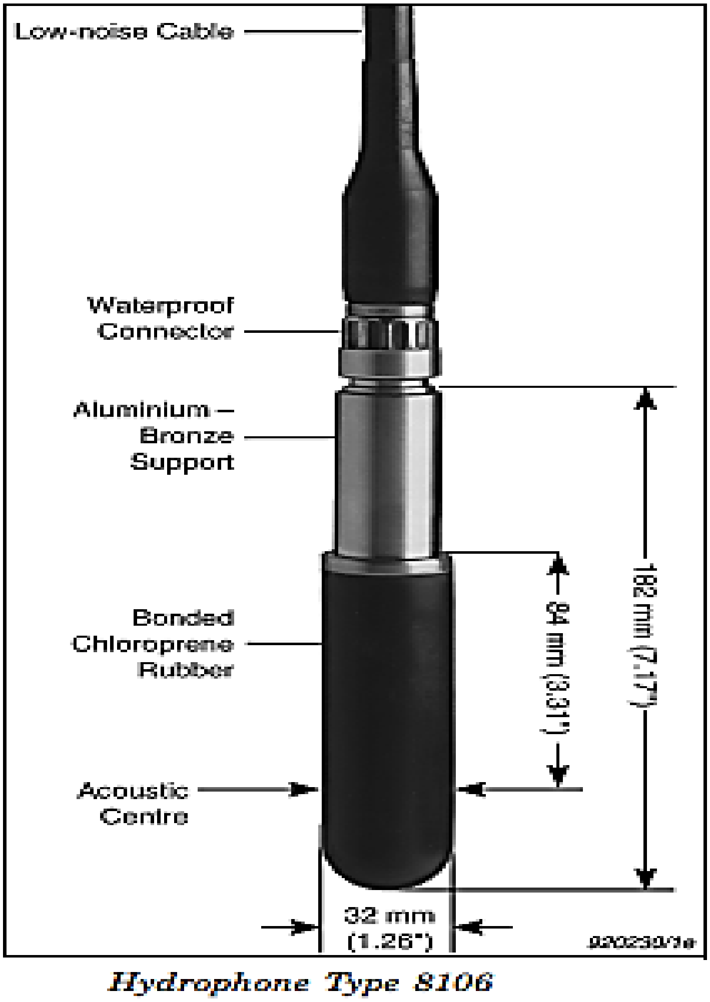
5.7.1 Acoustic Wave Sensors

Acoustic wave sensors are so named because their detection mechanism is a mechanical, or acoustic, wave. As the acoustic wave propagates through or on the surface of the material, any changes to the characteristics of the propagation path affect the velocity and/or amplitude of the wave. Changes in velocity can be monitored by measuring the frequency or phase characteristics of the sensor and can then be correlated to the corresponding physical quantity being measured. Virtually all acoustic wave devices and sensors use a piezoelectric material to generate the acoustic wave. Piezoelectricity refers to the production of electrical charges by the imposition of mechanical stress. The phenomenon is reciprocal. Applying an appropriate electrical field to a piezoelectric material creates a mechanical stress. Piezoelectric acoustic wave sensors apply an oscillating electric field to create a mechanical wave, which propagates through the substrate and is then converted back to an electric field for measurement. Among the piezoelectic substrate materials that can be used for acoustic wave sensors and devices, the most common are quartz (SiO_2), lithium tantalite ($LiTaO_3$), and, to a lesser degree, lithium niobate ($LiNbO_3$). An interesting property of quartz is that it is possible to select the temperature dependence of the material by the cut angle and the wave propagation direction. The advantage of using acoustic waves (vs electromagnetic waves) is the slow speed of propagation (5 orders of magnitude slower). For the same frequency, therefore, the wavelength of the elastic wave is 100,000 times shorter than the corresponding electromagnetic shortwave. This allows for the fabrication of verysmall sensors with frequencies into the gigahertz range with very fast response times. Solid state acoustic detectors have the electric circuit coupled to the mechanical structure where the waves propagate. The sensor generally has two (piezoelectric) transducers at each end. One at the transmitting end (generator) and one at the receiving end (receiver) where the wave is converted into an electric signal.

5.7.2Hydrophone

A **hydrophone** ('water + sound') is a microphone designed to be used underwater for recording or listening to underwater sound. Most hydrophones are based on a piezoelectric transducer that generates an electric potential when subjected to a pressure change, such as a sound wave. Some piezoelectric transducers can also serve as a sound projector, but not all have this capability, and some may be destroyed if used in such a manner.

A hydrophone (Figure 5.9) can detect airborne sounds, but will be insensitive because it is designed to match the acoustic impedance of water, a denser fluid than air. Sound travels 4.3 times faster in water than in air, and a sound wave in water exerts a pressure 60 times that exerted by a wave of the same amplitude in air. Similarly, a standard microphone can be buried in the ground, or immersed in water if it is put in a waterproof container, but will give poor performance due to the similarly bad acoustic impedance match.



Hydrophone Type 8106

Fig 5.9 Hydrophone

Sensors and Devices

How is sound measured?

Just as microphones are used to listen to sound in air, devices called hydrophones are used to listen to sound underwater. Microphones convert sound in air into electrical signals. The electrical signals can then be amplified, recorded, played back over loudspeakers, and transmitted over telephone lines. The electrical signals can also be used to measure the characteristics of the sound, such as amplitude and frequency. Similarly, hydrophones convert sound in water into electrical signals that can be amplified, recorded, played back over loudspeakers, and used to measure the characteristics of the sound. Hydrophones listen to sound, but do not transmit any sound (Fig 5.10).

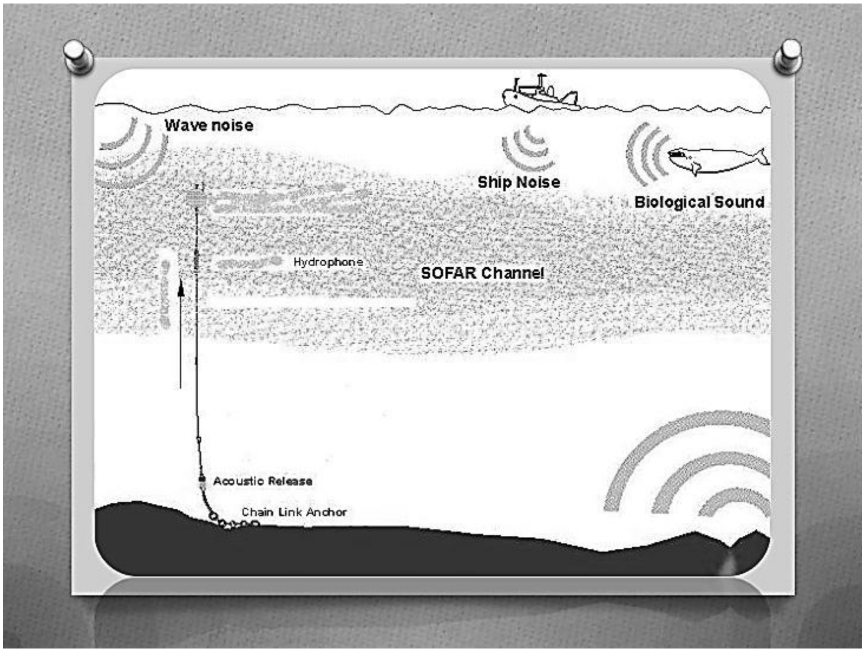


Fig 5.10 Diagram representation of Hydrophone

Most hydrophones are made from a piezoelectric material. This material has a special property that allows it to produce small electrical signals when squeezed, that is, when it is exposed to pressure changes. Because sound is a pressure wave, it can be detected by a piezoelectric element. Under the pressure of a sound wave, the piezoelectric element flexes and gives off electrical signals. These electrical signals can be recorded and later analyzed with computer programs to determine the properties of the sound wave, including amplitude and frequency.

Some hydrophones, called omnidirectional hydrophones, record sounds from all directions with equal sensitivity. Other hydrophones, called directional hydrophones, have a higher sensitivity to signals from a particular direction. Directional receivers are most often constructed using a number of omnidirectional hydrophones combined in what is called an array. Directional hydrophones are typically used in systems for locating and tracking objects.

5.8 LIGHT DEPENDENT RESISTOR (LDR)

A photoresistor or Light-Dependent Resistor (LDR) or photocell is a light-controlled variable resistor. The resistance of a photoresistor decreases with increasing incident light intensity; in other words, it exhibits photoconductivity as shown in Figure 5.11. A photoresistor can be applied in light-sensitive detector circuits, and light- and dark-activated switching circuits.

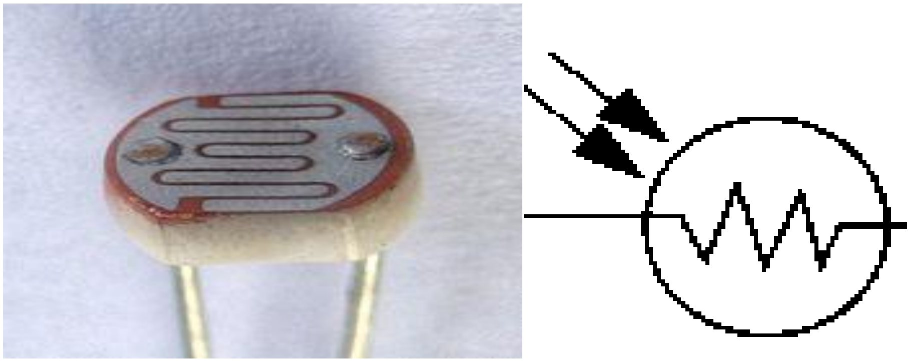


Fig.5.11 The symbol for a photoresistor

Working Principle of LDR

A light dependent resistor works on the principle of photo conductivity. Photo conductivity is an optical phenomenon in which the materials conductivity is increased when light is absorbed by the material. When light falls i.e. when the photons fall on the device, the electrons in the valence band of the semiconductor material are excited to the conduction band. These photons in the incident light should have energy greater than the band gap of the semiconductor material to make the electrons jump from the valence band to the conduction band. Hence when light having enough energy strikes on the device, more and more electrons are excited to the conduction band which results in large number of charge carriers. The result of this process is more and more current starts flowing through the device when the circuit is closed and hence it is said that the resistance of the device has been decreased. This is the most common working principle of LDR.

Characteristics of LDR

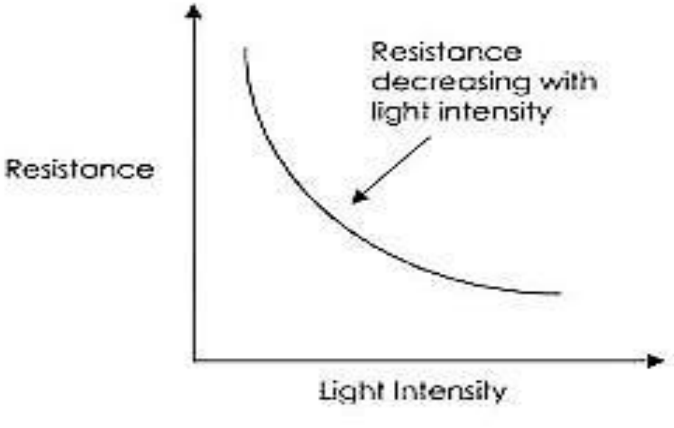


Fig.5.12 Variation of resistance Vs light intensity

LDR’s are light dependent devices whose resistance is decreased when light falls on them and that is increased in the dark. When a light dependent resistor is kept in dark, its resistance is very high. This resistance is called as dark resistance. It can be as high as 10^{12} Ω and if the device is allowed to absorb light its resistance will be decreased drastically. If a constant voltage is applied to it and intensity of light is increased the current starts increasing. The Figure 5.12 shows resistance Vs illumination curve for a particular LDR .

When light is incident on a photocell it usually takes about 8 to 12ms for the change in resistance to take place, while it takes one or more seconds for the resistance to rise back again to its initial value after removal of light. This phenomenon is called as resistance recovery rate. This property is used in audio compressors. Also, LDR’s are less sensitive than photo diodes and photo transistor.

Types of Light Dependent Resistors

Based on the materials used they are classified as:

i) Intrinsic photo resistors (Undoped semiconductor): These are made of pure semiconductor materials such as silicon or germanium. Electrons get excited from valance band to conduction band when photons of enough energy fall on it and number of charge carriers is increased.

ii) Extrinsic photo resistors: These are semiconductor materials doped with impurities which are called as dopants. These dopants create new energy bands above the valence band which are filled with electrons. Hence this reduces the band gap and less energy is required in exciting them. Extrinsic photo resistors are generally used for long wavelengths.

Construction of a Photocell

The structure of a light dependent resistor consists of a light sensitive material which is deposited on an insulating substrate such as ceramic. The material is deposited in zigzag pattern in order to obtain the desired resistance & power rating.

This zigzag area separates the metal deposited areas into two regions. Then the ohmic contacts are made on either sides of the area. The resistances of these contacts should be as less as possible to make sure that the resistance mainly changes due to the effect of light as shown in Figure 5.13. Materials normally used are cadmium sulphide, cadmium selenide, indium antimonide and cadmium sulphonide. The use of lead and cadmium is avoided as they are harmful to the environment.

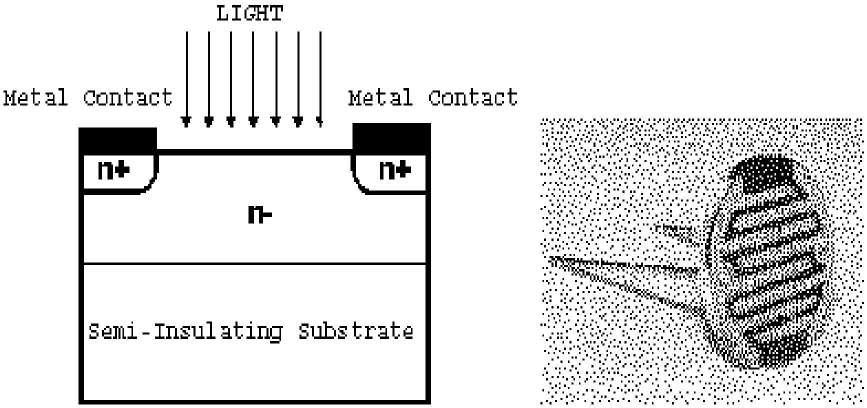


Fig.5.13 Construction of a Photocell

Applications of LDR

1. LDR’s have low cost and simple structure.
2. They are often used as light sensors.
3. They are used when there is a need to detect absence or presence of light as in a camera light meter.
4. Used in street lamps, alarm clock, burglar alarm circuits, light intensity meters, for counting the packages moving on a conveyor belt, etc.

5.9 PHOTODIODE

Introduction

A photodiode is a semiconductor device that converts light into current. The current is generated when photons are absorbed in the photodiode. A small amount of current is also produced when no light is present. Photodiodes may contain optical filters, built-in lenses, and may have large or small surface areas. Photodiodes usually have a slower response time as their surface area increases. The common, traditional solar cell used to generate electric solar power is a large area photodiode. Photodiodes are similar to regular semiconductor diodes except that they may be either exposed (to detect vacuum UV or X-rays) or packaged with a window or optical fiber connection to allow light to reach the sensitive part of the device. Many diodes designed for use specifically as a photodiode use a PIN junction rather than a p–n junction, to increase the speed of response. A photodiode is designed to operate in reverse bias.

Principle of Operation

A photodiode is a p–n junction or PIN structure. When a photon of sufficient energy strikes the diode, it creates an electron-hole pair. This mechanism is also known as the inner photoelectric effect. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in electric field of the depletion region. Thus holes move toward the anode, and electrons toward the cathode, and a photocurrent is produced. The total current through the photodiode is the sum of the dark current (current that is generated in the absence of light) and the photocurrent, so the dark current must be minimized to maximize the sensitivity of the device.

Photovoltaic Mode

When used in zero bias or photovoltaic mode, the flow of photocurrent out of the device is restricted and a voltage builds up. This mode exploits the photovoltaic effect, which is the basis for solar cells – a traditional solar cell is just a large area photodiode.

Photoconductive mode

In this mode the diode is often reverse biased (with the cathode driven positive with respect to the anode). This reduces the response time because the additional reverse bias increases the width of the depletion layer, which decreases the junction's capacitance. The reverse bias also increases the dark current without much change in the photocurrent. For a given spectral distribution, the photocurrent is linearly proportional to the illuminance (and to the irradiance). Although this mode is faster, the photoconductive mode tends to exhibit more electronic noise.

I-V Characteristic of a photodiode

The linear load lines represent the response of the external circuit and is shown in Figure 5.14. I= (Applied bias voltage-Diode voltage)/Total resistance. The points of intersection with the curves represent the actual current and voltage for a given bias, resistance and illumination.

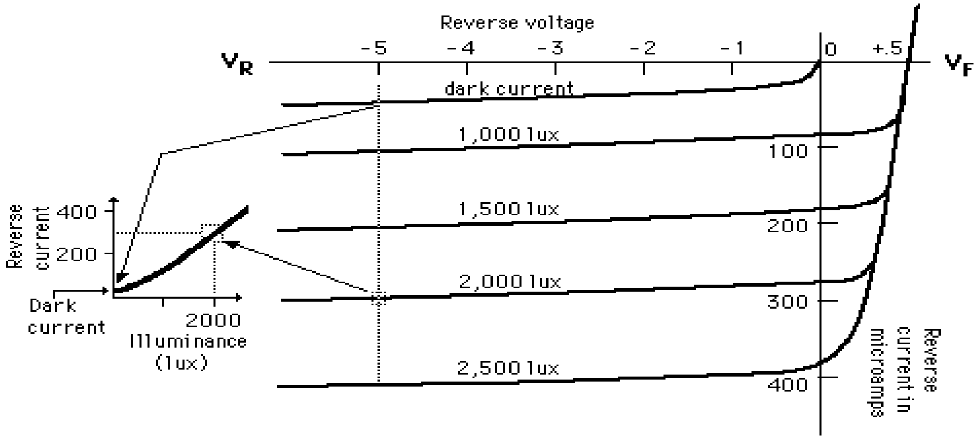


Fig 5.14 I-V Characteristic of a photodiode

Materials commonly used to produce photodiodes

The material used to make a photodiode is critical in defining its properties, because only photons with sufficient energy to excite electrons across the material's band gap will produce significant photocurrents.

Material	**Electromagnetic spectrumwavelength range (nm)**
Silicon	190–1100
Germanium	400–1700
Indium gallium arsenide	800–2600
Lead (II) sulfide	<1000–3500
Mercury cadmium telluride	400–14000

Applications of photodiode

1. p–n photodiodes are used in similar applications to other photodetectors, such as photoconductors, charge-coupled devices, and photomultipliers tubes.

2. Photodiodes are used in consumer electronics devices such as compact disc players, smoke detectors and the receivers for infrared remote-control devices used to control equipment from televisions to air conditioners. Either type of photosensor may be used for light measurement, as in camera light meters, or to respond to light levels, as in switching on street lighting after dark.

3. Photosensors of all types may be used to respond to incident light, or to a source of light which is part of the same circuit or system. A photodiode is often combined into a single component with an emitter of light, usually a light-

emitting diode (LED), either to detect the presence of a mechanical obstruction to the beam (slotted optical switch), or to couple two digital or analog circuits while maintaining extremely high electrical isolation between them, often for safety (optocoupler).

4. Photodiodes are often used for accurate measurement of light intensity in science and industry. They generally have a more linear response than photoconductors.

5. They are also widely used in various medical applications, such as detectors for computed tomography (coupled with scintillators), instruments to analyze samples (immunoassay), and pulse oximeters.

6. PIN diodes are much faster and more sensitive than p–n junction diodes and hence are often used for optical communications and in lighting regulation.

5.10 MAGNETIC SENSOR

Magnetic sensors detect changes and disturbances in a **magnetic** field like flux, strength and direction. Other types of detection **sensors** work with characteristics like temperature, pressure, light. The techniques used to create these **sensors** involve various combinations of physics and electronics.

Magnetic sensors can be split into two different types of sensor. This depends on whether they measure the vector components or the total magnetic field. While the results of each sensor vary, the techniques used to create these magnetic sensors use similar aspects of physics and electronic theories. All magnetic sensors work with the Earth's magnetic field.

The most common technologies used to create these magnetic sensors are the coil, fluxgate, optically pumped, nuclear precession, SQUID, Hall-effect, anisotropic magnetoresistance, giant magnetoresistance, magnetic tunnel junctions, giant magnetoimpedance, piezoelectric composites, magneto diode, magneto transistor, fiber optic, magneto optic as well as microelectromechanical systems-based magnetic sensors.

5.11 HALL EFFECT SENSOR

Hall effect sensor is the type of sensor which is typically used for detecting the practical behavior of anything such as motor, in response to changing magnetic field. Hall effect sensor is basically the transducers whose voltages varies by changing magnetic field. This sensor consists of a metal strip, when it is placed inside of any magnetic field then the presence of this magnetic field electron is deflected towards the edges of this strip then EMS is induced across its edges. In simple words it just works like as an analogue transducer. They are mainly used for speed detection, approximate switching, current and positioning sensing applications. A simple hall effect sensor is shown is Figure 5.15 which consists of three terminals.

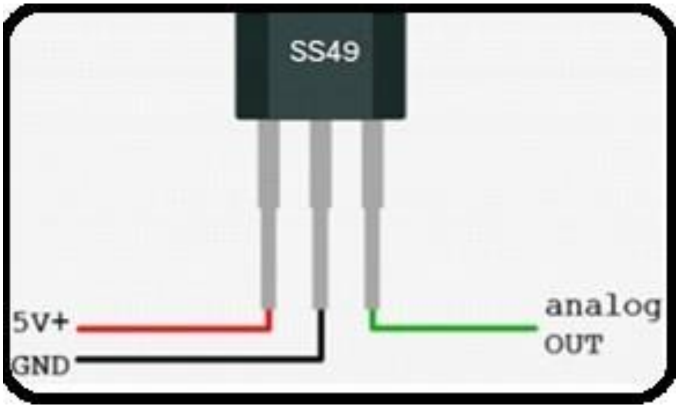


Fig.5.15 Hall Effect Sensor

Working Principle of Hall Effect Sensors

The working principle of hall effect sensor is very simple. It actually consists of a current carrying metal strip when this current carrying metal strip is placed inside any transverse magnetic field then EMF is developed across the edges of this current carrying metal strip. The magnitude of this developed EMF is depending upon the density of the flux and mobility of electron. This property of this current carrying conductor is called hall effect. This hall effect element is typically used for sensing current and magnetic field measurements. A hall effect element circuit is shown is Figure 5.16.

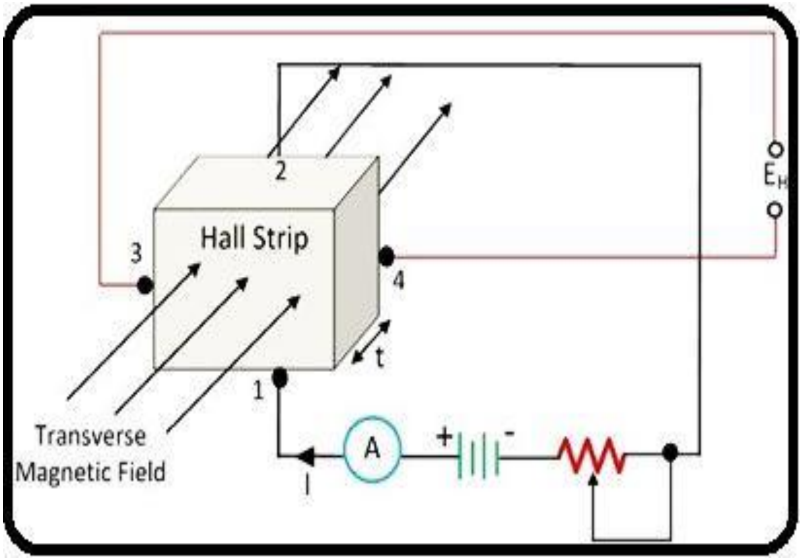


Fig.5.16 Hall Effect Sensor Circuit

According to the circuit when current is applied to terminal 1 and 3 then output is acquired at terminal 3 and 4. When no any magnetic field is applied to the strip then terminal 3 and 4 remains on same potential then there would be no any effect on hall effect strip. Similarly, when the magnetic field is applied on hall effect strip then EMF or voltages are induced across the terminal 3 and 4. These voltages are directly proportional to the material strength which have used for hall effect strip. These output voltages are obtained by the equation,

$$E_H = K_H IB/t \qquad (1)$$

Where K_H is hall effect co-efficient which is obtained from equation 2

$$K_H = v\text{-}m/A\text{-}wbm^{-2} \quad (2)$$

Where K_H is hall effect co-efficient, I is the current, B is the flux density in $Wbm^{-2}$and t is thickness of strip.

Through the output voltages, current and magnetic field strength could be easily measured but to measure the hall effect EMF in a simple conductor is very difficult because it is very small. Instead of simple conductor if we used the germanium conductor then its hall effect EMF could be easily measured with the help of moving coil instrument.

Different Applications of Hall Effect Sensor

Hall effect sensor has been used in different applications and its construction depends upon its configuration where it has been used. But these are mainly used in biomedical, automatic tellers, telecommunications, automobiles and process control industry.

- ***Position Sensor:*** This type of hall effect sensor is used of sensing sliding motion and in this sensor, there is tightly gap between hall effect element and magnet. When the magnet is moved in forward and backward direction then magnetic field is produced. The polarity of this magnetic field is positive when the element is moved toward north pole similarly the magnetic field is positive when the element is moved toward south pole. These sensors are called approximate sensor and used for position sensing.

- ***Brushless DC Motor Sensor:*** In brushless dc motors, power distribution is controlled through electronic commutation instead of mechanically commutation. For this purpose, three digital output hall effect sensors are installed at the stator end of brushless dc motor and for operating these sensors permanent magnet materials are installed at rotor shaft.

Sensors and Devices

- ***Current Sensor:*** It is the type of hall effect sensor which is used for measuring both ac and dc currents. These are available in range from 250mA to thousand amps. These are linear sensors and when this sensor is placed near magnetic field then voltage is developed across this sensor. The values of these voltages are proportional to magnetic field strength.

Principle and Structure

Here the magnetic sensor is based on hall effect. The effect is based on the interaction between moving electric carriers and an external magnetic field. In metal, these carriers are electrons. When an electron moves through a magnetic field, upon it acts a sideways force $F = qvB$, Where q is electronic charge, v is the speed of an electron and B is the magnetic field.

Hall Probe

A Hall probe contains an indium compound semiconductor crystal such as indium antimonide, mounted on an aluminum backing plate, and encapsulated in the probe head. The plane of the crystal is perpendicular to the probe handle. Connecting leads from the crystal are brought down through the handle to the circuit box.

When the Hall probe is held so that the magnetic field lines are passing at right angles through the sensor of the probe, the meter gives a reading of the value of magnetic flux density (B). A current is passed through the crystal which, when placed in a magnetic field has a "Hall effect" voltage developed across it. The Hall effect is seen when a conductor is passed through a uniform magnetic field. The natural electron drift of the charge carriers causes the magnetic field to apply a Lorentz force (the force exerted on a charged particle in an electromagnetic field) to these charge carriers. The result is what is seen as a charge separation, with a buildup of either positive or negative charges on the bottom or on the top of the plate. The crystal measures 5 mm square. The probe handle, being made of a non-ferrous material, has no disturbing effect on the field. A Hall probe should be calibrated against a known value of magnetic field strength. For a solenoid the Hall probe is placed in the center.

Working principle

When a beam of charged particles passes through a magnetic field, forces act on the particles and the beam is deflected from a straight path. The flow of electrons through a conductor is known as a beam of charged carriers. When a conductor is placed in a magnetic field perpendicular to the direction of the electrons, they will be deflected from a straight path. As a consequence, one plane of the conductor will become negatively charged and the opposite side will become positively charged. The voltage between these planes is called Hall voltage.

When the force on the charged particles from the electric field balances the force produced by magnetic field, the separation of them will stop. If the current is not changing, then the Hall voltage is a measure of the magnetic flux density. Basically, there are two kinds of hall effect sensors. One is linear which means the output of voltage linearly depends on magnetic flux density; the other is called threshold which means there will be a sharp decrease of output voltage at each magnetic flux density.

Materials for Hall effect sensors

The key factor determining sensitivity of Hall effect sensors is high electron mobility. As a result, the following materials are especially suitable for Hall effect sensors:

- gallium arsenide (GaAs)
- indium arsenide (InAs)
- indium phosphide (InP)
- indium antimonide (InSb)
- graphene

Signal processing and interface

Hall effect sensors are linear transducers. As a result, such sensors require a linear circuit for processing of the sensor's output signal. Such a linear circuit provides a constant driving current to the sensors and pro amplifies the output signal.

In some cases the linear circuit may cancel the offset voltage of Hall effect sensors. Moreover, AC modulation of the driving current may also reduce the influence of this offset voltage. Hall effect sensors with linear transducers are commonly integrated with digital electronics. This enables advanced corrections of the sensor's characteristics (e.g. temperature coefficient corrections) and digital interfacing to microprocessor systems. In some solutions of IC hall effect sensors a DSP is used, which provides for more choices among processing techniques.

The Hall effect sensor interfaces may include input diagnostics, fault protection for transient conditions, and short/open circuit detection. It may also provide and monitor the current to the Hall effect sensor itself. There are precision IC products available to handle these features.

Advantages

1. A Hall effect sensor may operate as an electronic switch. Such a switch costs less than a mechanical switch and is much more reliable.
2. It can be operated up to 100 kHz.
3. It does not suffer from contact bounce because a solid state switch with hysteresis is used rather than a mechanical contact.
4. It will not be affected by environmental contaminants since the sensor is in a sealed package Therefore, it can be used under severe conditions.
5. In the case of linear sensor (for the magnetic field strength measurements), a Hall effect sensor: can measure a wide range of magnetic fields is available that can measure either North or South pole magnetic fields can be flat.

Disadvantages

Hall effect sensors provide much lower measuring accuracy than fluxgate magnetometers or magnetoresistance-based sensors. Moreover, Hall effect sensors drift significantly, requiring compensation.

Applications

Position sensing

Sensing the presence of magnetic objects (connected with the position sensing) is the most common industrial application of Hall effect sensors, especially those operating in the switch mode (on/off mode). The Hall effect sensors are also used in the brushless DC motor to sense the position of the rotor and to switch the transistors in the right sequence.

Smartphones use hall sensors to determine if the Flip Cover accessory is closed.

Direct current (DC) transformers

Hall effect sensors may be utilized for contactless measurements of DC current in current transformers. In such a case the Hall effect sensor is mounted in the gap in magnetic core around the current conductor. As a result, the DC magnetic flux can be measured, and the DC current in the conductor can be calculated.

Automotive fuel level indicator

The Hall sensor is used in some automotive fuel level indicators. The main principle of operation of such indicator is position sensing of a floating element. This can either be done by using a vertical float magnet or a rotating lever sensor.

In a vertical float system a permanent magnet is mounted on the surface of a floating object. The current carrying conductor is fixed on the top of the tank lining up with the magnet. When the level of fuel rises, an increasing magnetic field is applied on the current resulting in higher Hall voltage. As the fuel level decreases, the Hall voltage will also decrease. The fuel level is indicated and displayed by proper signal condition of Hall voltage. In a rotating lever sensor a diametrically magnetized ring magnet rotates about a linear hall sensor. The sensor only measures the perpendicular (vertical) component of the field. The strength of the field measured correlates directly to the angle of the lever and thus the level of the fuel tank.

Keyboard switch

Developed by Everett A. Vorthmann and Joeseph T. Maupin for Micro Switch (a division of Honeywell) in 1969, the switch was known to still be in production until as late as 1990. The switch is one of the highest quality keyboard switches ever produced, with reliability being the main aim of the design. The key-switches have been tested to have a lifetime of over 30 billion keypresses, the switch also has dual open-collector outputs for reliability. The Honeywell Hall Effect switch is most famous used in the Space-cadet keyboard, a keyboard used on LISP machines.

5.12 STRAIN GAUGES

A strain gauge (or strain gage) is a device used to measure strain on an object. Invented by Edward E. Simmons and Arthur C. Ruge in 1938, the most common type of strain gauge consists of an insulating flexible backing which supports a metallic foil pattern. The gauge is attached to the object by a suitable adhesive, such as cyanoacrylate. As the object is deformed, the foil is deformed, causing its electrical resistance to change. This resistance change, usually measured using a Wheatstone bridge, is related to the strain by the quantity known as the gauge factor. Strain gauges are devices whose resistance changes under the application of force or strain. They can be used for measurement of force, strain, stress, pressure, displacement, acceleration etc.

Sensors and Devices

Principle of Strain Gauges

Each metal has its specific resistance. An external tensile force / (compressive force) increases/decreases the resistance by elongating/contracting it. Suppose the original resistance is R and a strain-initiated change in resistance is 'ΔR'.

$$\text{Gauge factor (K)} = \frac{\Delta R / R}{\Delta l / l}$$

where, 'K' is a gauge factor, the coefficient expressing strain gauge sensitivity. General purpose strain gauges use copper-nickel or nickel-chrome alloy for the resistive element.

Derivation of Strain Gauges

The Strain Gauge Derivation is an example of a passive transducer that uses the variation in electrical resistance in wires to sense the strain produced by a force on the wires. It is well known that stress (force/unit area) and strain (elongation or compression/unit length) in a member or portion of any object under pressure is directly related to the modulus of elasticity. Since strain can be measured more easily by using variable resistance transducers, it is a common practice to measure strain instead of stress, to serve as an index of pressure. Such transducers are popularly known as strain gauges

When force is applied to any metallic wire its length increases due to the strain.

Thus, a tensile stress tends to elongate the wire and thereby increase its length and decrease its cross-sectional area. The combined effect is an increase in resistance, as seen from the following equation

$$R = \rho \times l/A$$

where 'ρ' = the specific resistance of the material in Ωm.

‘l’ = the length of the conductor in m

‘A’ = the area of the conductor in m^2

As a result of strain, two physical parameters are of particular interest. The measurement of the sensitivity of a material to strain is called the gauge factor (GF).It is the ratio of the change in resistance ΔR/R to the change in the length Δl/l

$$GF(K) = \frac{\Delta R / R}{\Delta l / l}$$

where ‘K’ = gauge factor

‘ΔR’ = the change in the initial resistance in Ω’s

‘R’ = the initial resistance in Ω (without strain)

‘Δl’ = the change in the length in m

‘l’ = the initial length in m (without strain)

Since, strain is defined as the change in length divided by the original length,

$$\text{i.e. } \sigma = \frac{\Delta l}{l} \qquad (1)$$

Eqn (1) can be written as

$$K = \frac{\Delta R / R}{\sigma} \qquad (2)$$

where ‘σ’ is the strain in the lateral direction. The resistance of a conductor of uniform cross-section is

$$R = \rho \frac{\text{length}}{\text{area}}$$

$$R = \rho \frac{l}{\pi r^2}$$

$$r = \frac{d}{r} \therefore r^2 = \frac{d^2}{4}$$

$$R = \rho \frac{1}{\pi d^2 / 4} = \rho \frac{1}{\pi / 4d^2} \qquad (3)$$

where ‘ρ’ = specific resistance of the conductor

‘l’ = length of conductor

‘d’ = diameter of conductor

When the conductor is stressed, due to the strain, the length of the conductor increases by ‘Δl’ and the simultaneously decreases by ‘Δd’ in its diameter. Hence the resistance of the conductor can now be written as,

$$R_s = \rho \frac{(1 + \Delta l)}{\pi / 4(d - \Delta d)^2} = \frac{\rho(1 + \Delta l)}{\pi / 4(d^2 - 2d\Delta d + \Delta d^2)}$$

Since ‘Δd’ is small, Δd^2 can be neglected,

$$R_s = \rho \frac{(1 + \Delta l)}{\pi / 4(d^2 - 2d\Delta d)}$$

$$= \frac{\rho(1 + \Delta l)}{\pi / 4d^2 \left(1 - \frac{2\Delta d}{d}\right)} = \frac{\rho l(1 + \Delta l / l)}{\pi / 4d^2 \left(1 - \frac{2\Delta^2}{d}\right)} \qquad (4)$$

Now, Poisson’s ratio ‘μ’ is defined as the ratio of strain in the lateral direction to strain in the axial direction, that is,

$$R_s = \frac{\rho l(1+\Delta l/l)}{(\pi/4)d^2(1-2\mu\Delta l/l)} \qquad (5)$$

Rationalizing we get.

$$R_s = \frac{\rho l(1+\Delta l/l)}{(\pi/4)d^2(1-2\mu\Delta l/l)}\frac{(1+2\mu\Delta l/l)}{(1+2\mu\Delta l/l)}$$

$$R_s = \frac{\rho l}{(\pi/4)d^2}\left[\frac{(1+\Delta l/l)}{(1+2\mu\Delta l/l)}\frac{(1+2\mu\Delta l/l)}{(1+2\mu\Delta l/l)}\right]$$

$$R_s = \frac{\rho l}{(\pi/4)d^2}\left[\frac{1+2\mu\Delta l/l+2\Delta l/l+2\mu\Delta l/le\Delta l/l}{1-4\mu^2(\Delta l/l)^2}\right]$$

$$R_s = \frac{\rho l}{(\pi/4)d^2}\left[\frac{1+2\mu\Delta l/l+\delta l/l+2\mu\Delta l^2/l^2}{1-4\mu^2\Delta l^2/l^2)}\right] \qquad (6)$$

Since ‘Δl’ is small, we can neglect higher powers of ‘Δl’,

$$R_s = \frac{\rho l}{(\pi/4)d^2}$$

$R_s = R + \Delta R$

$$\Delta R = \frac{\rho l}{(\pi/4)d^2}(\Delta l/l)(1+2\mu) \qquad (7)$$

The gauge factor will now be

$$K = \frac{\Delta R / R}{\Delta l / l} = \frac{(\Delta l / l)(1 + 2\mu)}{\Delta l / l}$$

$$= 1 + 2\mu$$

$$\mathbf{K = 1 + 2\mu}$$

Types of strain gauges

The following types of strain gauges are the most important.

1. Wire strain gauges
2. Foil strain gauges
3. Semiconductor strain gauges

Resistance Wire Gauge

Resistance wire gauges are used in two basic forms, the unbonded type, and the bonded type.

1.Unbonded Resistance Wire Strain Gauge

An unbonded strain gauge consists of a wire streched between two points in an insulating medium, such as air. The diameter of the wire used is about 25 µm. The wires are kept under tension so that there is no sag and no free vibration. Unbonded Strain Gauge Derivation are usually connected in a bridge circuit. The bridge is balanced with no load applied as shown in Fig. 5.17. When an external load is applied, the resistance of the Strain Gauge Derivation changes, causing an unbalance of the bridge circuit resulting in an output voltage. This voltage is proportional to the strain. A displacement of the order of 50µm can be detected with these strain gauges.

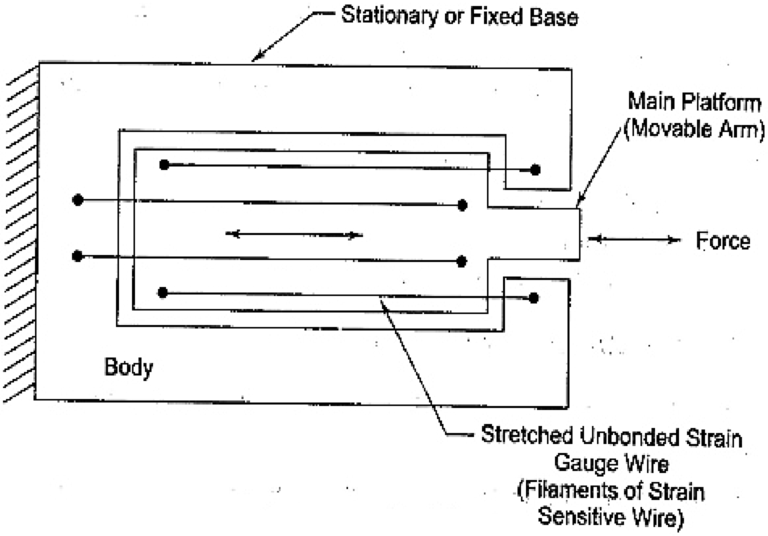


Figure 5.17 Unbounded Strain Gauge

Bonded Resistance Wire Strain Gauges

A metallic bonded Strain Gauge Derivation is shown in Figure 5.18.

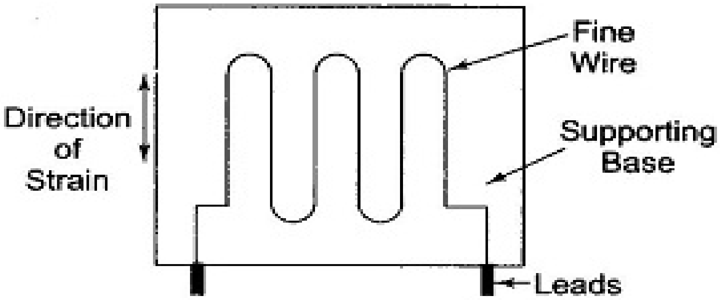


Figure 5.18 Bonded resistance wire Strain Gauge

Applications of the Strain Gauges

The strain gauges are used for two main purposes:

1) Measurement of strain: Whenever any material is subjected to high loads, they come under strain, which can be measured easily with the strain gauges. The strain can also be used to carry out stress analysis of the member.

2) Measurement of other quantities: The principle of change in resistance due to applied force can also be calibrated to measure a number of other quantities like force, pressure, displacement, acceleration etc since all these parameters are related to each other. The strain gauges can sense the displacements as small as 5 µm. They are usually connected to the mechanical transducers like bellows for measuring pressure and displacement and other quantities

Warren Seymour Johnson

Warren Seymour Johnson (November 6, 1847 – December 5, 1911) was an American college professor who was frustrated by his inability to regulate individual classroom temperatures. His multi-zone pneumatic control system solved the problem. Johnson's system for temperature regulation was adopted worldwide for office buildings, schools, hospitals, and hotels – essentially any large building with multiple rooms that required temperature regulation. To manufacture and market his system, Johnson established the Johnson Electric Service Company which eventually became Johnson Controls.

Johnson had an inquisitive mind and was particularly interested in electricity. In 1883, he developed a thermostat, which he deployed at the State Normal School. He called the instrument an "electric tele-thermoscope" in the patent application. It was a bi-metal coiled thermostat with a mercury switch, which could be used to ring a bell to alert the fireman to open or close the heating damper. While not the first bi-metal thermostat, Johnson received a patent for the device and interested William Plankinton, heir to the Plankinton Packing Company, to provide financial backing to manufacture the device.

While it might have seemed crude by the modern standards that we have today, this thermostat was able to keep temperatures within a degree of accuracy something that a better than some of the low quality thermostats on the market today! The first motion sensor used for an alarm system came about in the early part of the 1950s, and was the invention of Samuel Bagno. His device made use of ultrasonic frequencies as well as the Doppler Effect

QUESTIONS

1. Briefly explain the principle, construction and working of Bourdon pressure.
2. Briefly explain the principle, construction and working of vibration sensor.
3. Define acoustic sensor and explain the principle, types.
4. Illustrate with suitable diagram, the principle and working of LDR.
5. Explain in brief the principle and working of temperature sensor. Give its advantages and disadvantages.
6. Write notes on magnetic sensor-Hall effect sensors.
7. Define Strain Gauge. Explain the principle, types and applications of strain gauge.
8. What is a photo diode? Explain the two modes of its operations.