HTENG 113 PHYSICS FOR ENGINEERS NOTES

COMPILED BY V E GORA

ELECTRONIC DEVICES

ATOMIC STRUCTURE

Electron shells and orbits

Electrons orbit the nucleus of an atom at certain distances from the nucleus. Each orbit from the nucleus corresponds to a certain energy level. In an atom the orbits are grouped into energy bands known as shells designated by letters K, L, M, N…..etc with K being closest to the nucleus.

Valence electrons

The outermost shell is known as the valence shell and the electrons in this orbit are called valence electrons. These valence electrons contribute to chemical reactions and bonding within the structure of a material and determine its electrical properties.

The number of electrons in each shell

The number of electrons (N_e) that can exist in each shell of an atom is determined by the formula $N_e = 2n^2$ where n is the number of the shell. The innermost (K) shell is number 1, L is 2, N is 3 etc. According to the above formula the maximum number of electrons in K, L, M and N shells are 2, 8, 18 and 32 respectively. All shells in a given atom must be completely filled with electrons except the outer (valence) shell.

SEMICONDUCTORS, CONDUCTORS AND INSULATORS

Conductors

A conductor is a material that easily conducts an electric current eg copper, aluminium and silver etc.

Insulators

An insulator is a material that does not conduct electrical current under normal circumstances eg paper, wood air, plastic etc..

Semiconductors

A semiconductor is a material that is between conductors and insulators in its ability to conduct an electrical current eg silicon, carbon and germanium etc. The single-element semiconductors are characterized by atoms with four valence electrons. Compound semiconductors such as gallium arsenide are also commonly used.
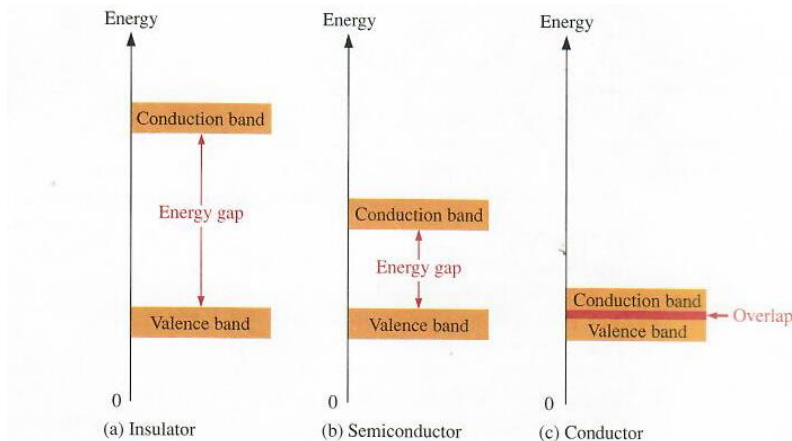
Energy bands

FIGURE 1–5
Energy diagrams for the three types of materials.

If an electron acquires enough additional energy from an external source, it leaves the valence shell and becomes a free electron and exists in what is known as the conduction band. The difference in energy between the valence band and the conduction band is called the energy gap. This is the amount of energy that a valence electron must have in order to jump from the valence band to the conduction band.

Fig 1-5 shows energy diagrams for three materials:

  (a) Insulators have a very wide energy gap. Valence electrons do not jump into the conduction band except under breakdown conditions where extremely high voltages are applied across the material.
  (b) Semiconductors have a much narrower energy gap. This gap permits some valence electrons to jump into the conduction band and become free electrons.
  (c) The energy bands in conductors overlap. In a conductive material there is always a large number of free electrons.

Covalent bonds

Fig 1-8 shows how each silicon atom positions itself with four adjacent atoms to form a silicon crystal. A silicon atom with its four valence electrons shares an electron with each of its four neighbours. This effectively creates eight valence electrons for each atom and produces a state of chemical stability. This sharing of valence electrons produces the covalent bonds that hold the atoms together. An intrinsic crystal is one that has no impurities.

FIGURE 1–8
*Covalent bonds in
silicon.*
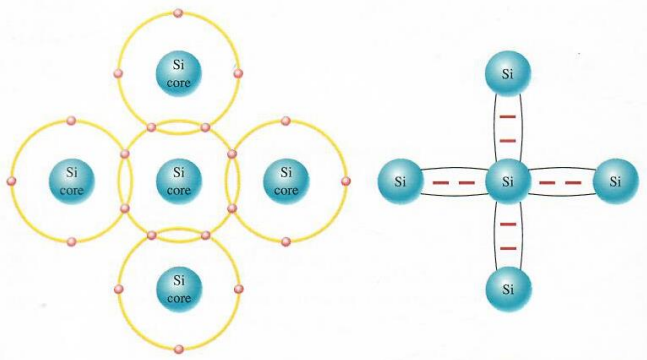
CONDUCTION IN SEMICONDUCTORS

Electrons of an atom can exist only within prescribed energy bands. Each shell around the nucleus corresponds to a certain energy band and is separated from adjacent shells by energy gaps, in which no electrons can exist. Fig 1-10 shows an unexcited (no external energy such as heat) silicon atom. This condition occurs only at a temperature of absolute $0°$ K.
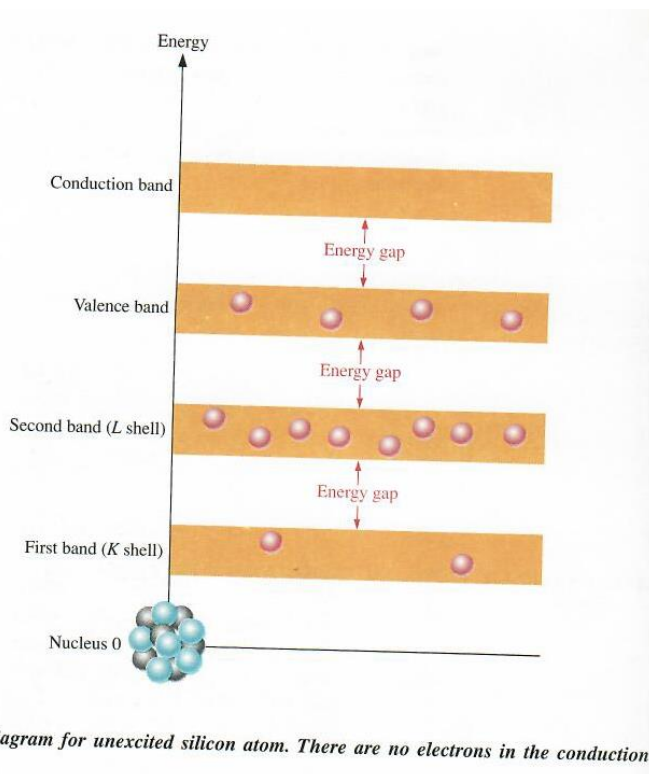

FIGURE 1–10
Energy band diagram for unexcited silicon atom. There are no electrons in the conduction band.

4

Conduction electrons and holes

An intrinsic (pure) silicon crystal at room temperature derives heat (thermal) energy from the surrounding air, causing some valence electrons to gain sufficient energy to jump the gap from the valence band into the conduction band, becoming free electrons not bound to any one atom but free to drift as conduction electrons as in Fig 1-11.

When an electron jumps to the conduction band, a vacancy is left in the valence band. This vacancy is called a hole. For every electron raised to the conduction band by external energy, there is one hole left in the valence band, creating what is known as an electron-hole pair. Recombination occurs when a conduction-band electron loses energy and falls back into a hole in the valence band.

Electron and hole current

When a voltage is applied across a piece of intrinsic silicon(Fig 1-3) ,the thermally generated free electrons in the conduction band, which are free to move randomly in the crystal structure, are now easily attracted to the positive end constituting one type of current in semiconductor material called an electron current. Another type of current occurs at the valence level, where the holes created by the free electrons exist. Electrons remaining in the valence band are still attached to their atoms and are not free to move randomly in the crystal structure as are the free electrons. However, a valence electron can move into a nearby hole, with little change in its energy level, thus leaving another hole where it came from. Effectively the hole has moved from one place to another in the crystal structure as a hole current(Fig 1-14).
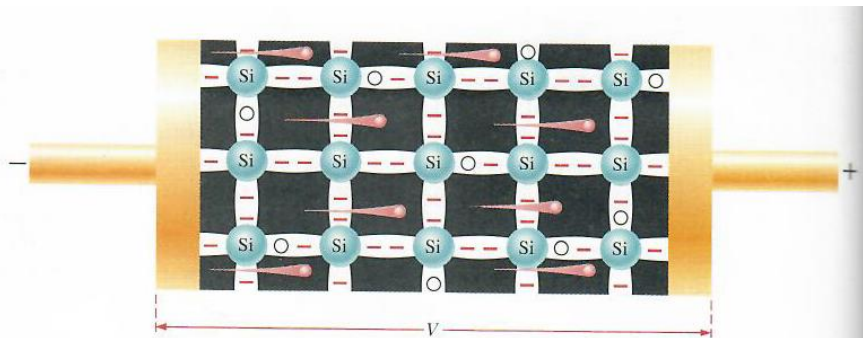

FIGURE 1–13
Electron current in intrinsic silicon is produced by the movement of thermally generated free electrons.
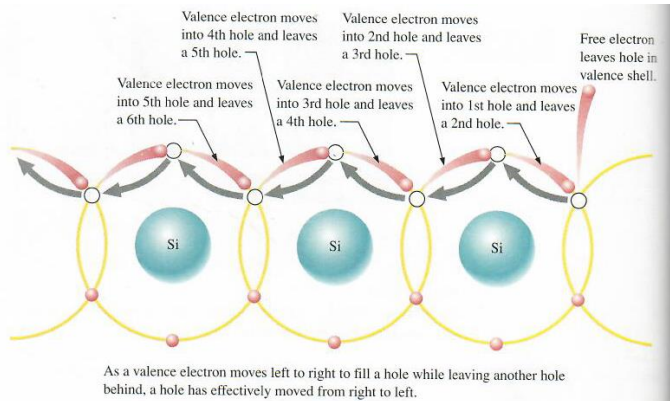
Valence electron moves into 4th hole and leaves a 5th hole.

Valence electron moves into 2nd hole and leaves a 3rd hole.

Free electron leaves hole in valence shell.

Valence electron moves into 5th hole and leaves a 6th hole.

Valence electron moves into 3rd hole and leaves a 4th hole.

Valence electron moves into 1st hole and leaves a 2nd hole.

As a valence electron moves left to right to fill a hole while leaving another hole behind, a hole has effectively moved from right to left.

FIGURE 1–14
Hole current in intrinsic silicon.

N-TYPE AND P-TYPE SEMICONDUCTORS

Doping

Doping is the controlled addition of impurities to the intrinsic (pure) semiconductor material to increase its conductivity. Doping increases the number of charge carriers (electrons or holes). The two categories of impurities are n-type and p-type.

N-Type semiconductor
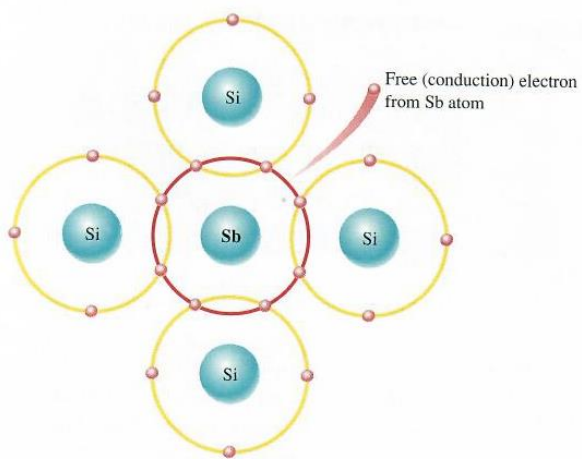

Free (conduction) electron from Sb atom

FIGURE 1–15
Pentavalent impurity atom in a silicon crystal structure. An antimony (Sb) impurity atom is shown in the center. The extra electron from the Sb atom becomes a free electron.

To increase the number of conduction-band electrons in intrinsic silicon, pentavalent impurity atoms are added eg arsenic(As), phosphorus(P), bismuth(Bi), antimony(Sb). As in

Fig 1-15 each pentavalent atom eg Sb forms covalent bonds with four adjacent silicon atoms, leaving one extra electron. This extra electron becomes a conduction electron because it is not attached to any atom. The pentavalent atoms are called donor atoms. In n-type semiconductors electrons are the majority carriers while holes are the minority carriers.

P-Type semiconductor

FIGURE 1–16
Trivalent impurity atom in a silicon crystal structure. A boron (B) impurity atom is shown in the center.
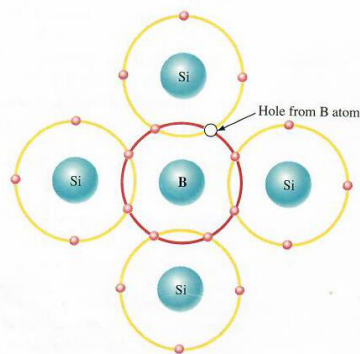
To increase the number of holes in intrinsic silicon, trivalent impurity atoms are added eg aluminium(Al), indium(In), gallium(Ga) and boron(B). As in Fig 1-16 each trivalent atom(boron) forms covalent bonds with four adjacent silicon atoms. All three of the boron's valence electrons are used in the covalent bonds, and since four electrons are required a hole results when each trivalent atom is added. Trivalent atoms are called acceptor qtoms. In p-type material holes are the majority carriers while electrons are the minority carriers.
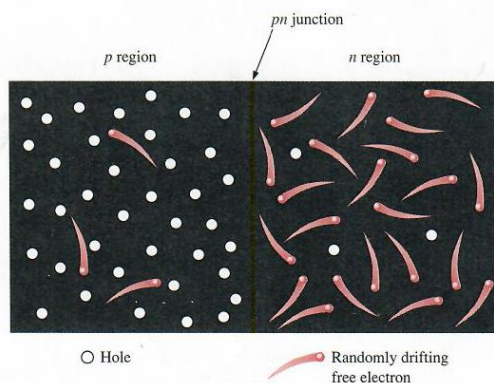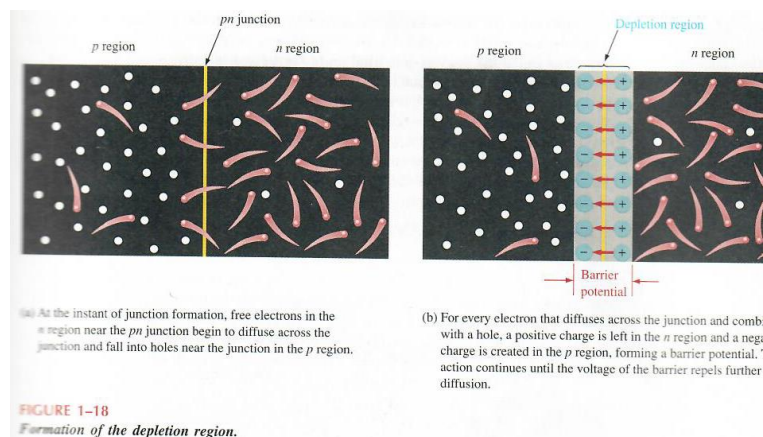
THE PN JUNCTION

FIGURE 1–17
The basic pn structure at the instant of junction formation showing only the majority and minority carriers.

If a piece of p-type silicon material is joined with a piece of n-type silicon material a pn junction is formed between the two regions(see Fig 1-17). The p-region has many holes (majority carriers) from the impurity atoms and only a few thermally generated free electrons (minority carriers). The n-region has many free electrons (majority carriers) from the impurity atoms and only a few thermally generated free holes (minority carriers).

Formation of the depletion region

(a) At the instant of junction formation, free electrons in the n region near the *pn* junction begin to diffuse across the junction and fall into holes near the junction in the *p* region.

(b) For every electron that diffuses across the junction and combines with a hole, a positive charge is left in the *n* region and a negative charge is created in the *p* region, forming a barrier potential. This action continues until the voltage of the barrier repels further diffusion.

FIGURE 1–18
Formation of the depletion region.

Before the junction is formed the free electrons in the n-region are randomly drifting in all directions. At the instant of the pn junction formation, the free electrons near the junction in the n-region begin to diffuse across the junction into the p-region where they combine with holes near the junction(Fig 1-18a).

When the pn junction is formed, the n-region loses free electrons as they diffuse across the junction. This creates a layer of positive charges (pentavalent ions) near the junction. As the electrons move across the junction, the p-region loses holes as the electrons and holes combine. This creates a layer of negative charges (trivalent ions) near the junction. These two layers of positive and negative charges form the depletion region(Fig 1-18b). The term depletion refers to the fact that the region near the pn junction is depleted of charge carriers (electrons and holes) due to diffusion across the junction.

As electrons continue to diffuse across the junction, more and more positive and negative charges are created near the junction as the depletion region is formed. A point is reached where the total charge in the depletion region repels ant further diffusion of electrons (negatively charged particles) into the p region (like charges repel) and the diffusion stops. The depletion region acts as a barrier to the further movement of electrons across the junction.

Barrier potential

In the depletion region there are many positive charges and many negative charges on opposite sides of the pn junction. The forces between the opposite charges form an electric

field as shown by arrows between the positive and negative charges. This electric field is a barrier to the free electrons in the n region. External energy must be applied to get the electrons to move across the barrier of the electric field in the depletion region. The potential difference of the electric field is called the barrier potential and it is 0.7V and 0.3V for silicon and germanium respectively.

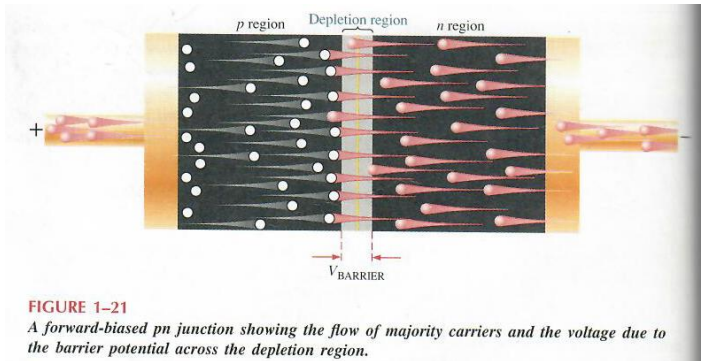**BIASING THE PN JUNCTION**

Forward bias

FIGURE 1–20
A pn junction connected for forward bias.
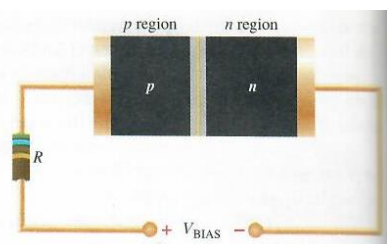
To bias a pn junction, apply an external dc voltage across it. This external bias voltage is designated as V_{BIAS}.(Fig 1-20) The resistance limits the current to a value that will not damage the pn structure. The negative side of V_{BIAS} is connected to the n region of the pn junction and the positive side is connected to the p region.(Fig 1-20) V_{BIAS} must be greater than the barrier potential. The negative side of the bias voltage pushes the free electrons which are the majority carriers in the n region towards the pn junction. This flow of free electrons is called the electron current. The negative side of the source also provides a continuous flow of electrons through the external connection(conductor) and into the n region as shown jn Fig 1-21. The bias-voltage imparts sufficient energy to the free electrons for them to overcome the barrier potential of the depletion region and move through into the p region. Once in the p region these conduction electrons have lost enough energy to immediately combine with holes in the valence band.
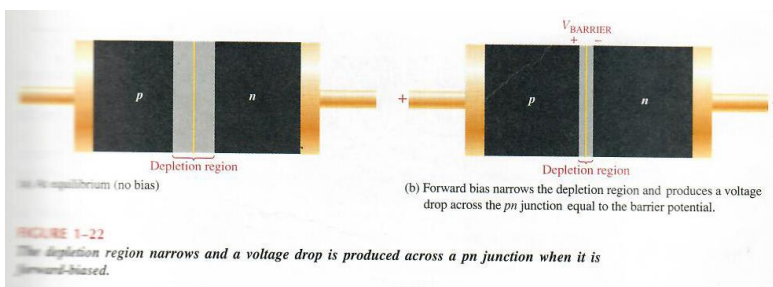
Now, the electrons are in the valence band in the p region, simply because they have lost too much energy overcoming the barrier potential to remain in the conduction band. The positive side of the bias-voltage source attracts the valence electrons toward the left end of the p region. The holes in the p region provide the medium for these valence electrons to move through the p region. The electrons move from one hole to the next towards the left. The holes in the p region which are the majority carriers in the p region effectively (not actually) move to the right as the hole current.

9

FIGURE 1–21
A forward-biased pn junction showing the flow of majority carriers and the voltage due to the barrier potential across the depletion region.

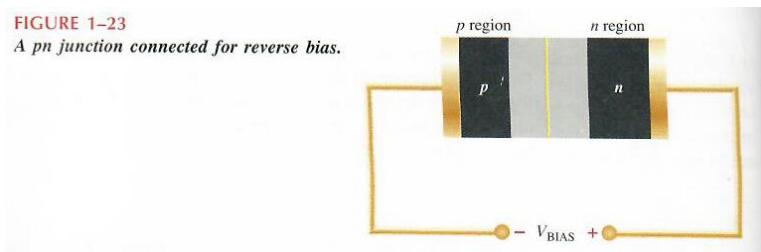**The effect of forward bias in the depletion region**

As more electrons flow into the depletion region, the number of positive ions is reduced. As more holes effectively flow into the depletion region the number of negative ions is reduced. The reduction of positive and negative ions during forward bias causes the depletion region to narrow.(Fig 1-22)

Effect of barrier potential during forward bias

FIGURE 1–22
The depletion region narrows and a voltage drop is produced across a pn junction when it is forward-biased.

When forward bias is applied, the free electrons are provided with enough energy from the bias voltage source to overcome the barrier potential and effectively "climb" the energy hill and cross the depletion region. This energy is equal to the barrier potential. The energy loss is equal to the barrier potential of 0.7V for silicon and 0.3V for germanium.

Reverse bias

FIGURE 1–23
A pn junction connected for reverse bias.

The positive side of V_{BIAS} is connected to the n region of the pn junction and the negative side is connected to the p region. The depletion region is much wider during reverse bias as compared to forward bias or at equilibrium.

In the n region as the electrons flow towards the positive side of the voltage source additional positive ions are created. This results in the widening of the depletion region and the depletion of majority carriers.

In the p region, electrons from the negative side of the voltage source enter as valence electrons and move from hole to hole toward the depletion region where they create additional negative ions. This results in the widening of the depletion region and the depletion of majority carriers.(Fig 1-24)
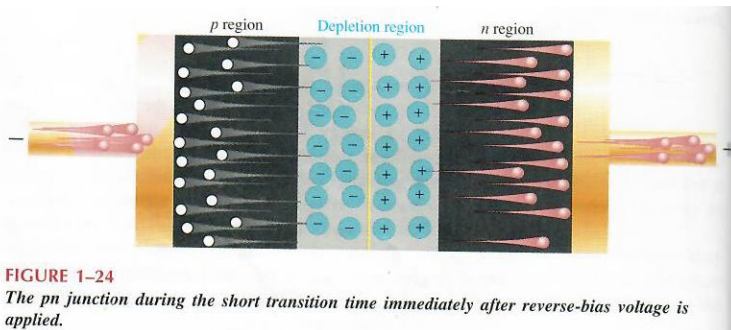

FIGURE 1–24
The pn junction during the short transition time immediately after reverse-bias voltage is applied.

Reverse current

The reverse current is extremely small and is caused by the minority carriers in the n and p regions that are produced by thermally generated electron-hole pairs.

Reverse breakdown

Normally the reverse current is so small that it can be neglected but however if the external reverse-bias voltage is increased to a value called the breakdown voltage, the reverse current will drastically increase. The high reverse-bias voltage imparts energy to the free minority electrons so that as they speed through the p region, they collide with atoms with enough energy to knock valence electrons out of orbit and into the conduction band. The newly created conduction electrons are also high in energy and repeat the process. As these high-energy electrons go through the depletion region, they have enough energy to go through the n region as conduction electrons, rather than combining with holes. This multiplication of conduction electrons is called an avalanche and results in a very high reverse current that can damage the pn structure because of excessive heat dissipation.

CURRENT-VOLTAGE CHARACTERISTIC OF A PN JUNCTION(DIODE)

With 0 V across the pn junction, there is no forward current I_F but as you gradually increase the bias voltage, the forward current and the voltage across the pn junction gradually increase. When the applied bias voltage is increased to a value where the voltage across the pn junction reaches approximately 0.7V(barrier potential), the forward current begins to increase rapidly.

When a reverse-bias voltage is applied across a pn junction, there is only an extremely small reverse current(I_R) through the junction(Fig 1-31). With 0 V across the pn junction there is no reverse current. As the reverse-bias voltage is gradually increased, a very small reverse current and the voltage across the pn junction also increases. When the applied bias voltage is increased to a value where the reverse voltage across the pn junction(V_R) reaches the breakdown value (V_{BR}), the reverse current begins to increase rapidly while the voltage above V_{BR} increases very little.
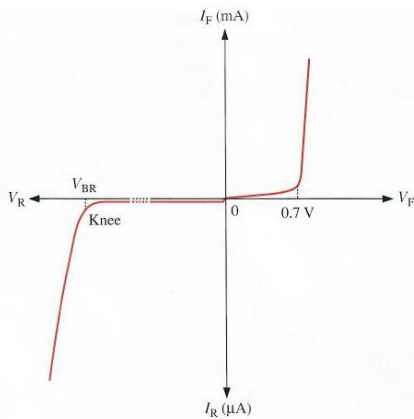
FIGURE 1–31
The complete I-V characteristic curve.

THE DIODE

The rectifier diode is a single pn junction device with conductive contacts and wire leads connected to each region(Fig 1-33). One half of the diode is an n-type semiconductor and the other half is a p-type semiconductor. The n region is called the cathode and the p region is called the anode. The arrow in the symbol points in the direction of conventional current(opposite to electron flow).
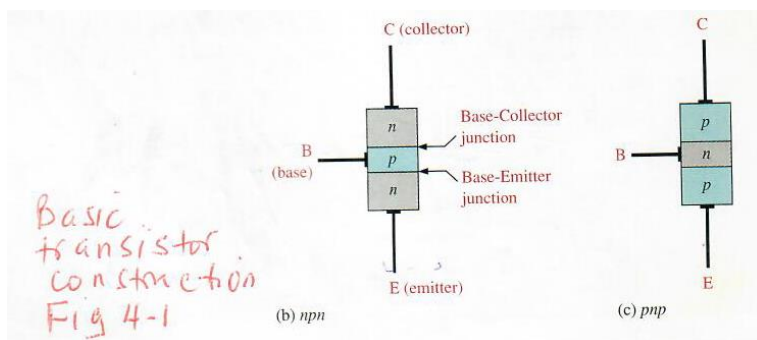
FIGURE 1-33
Diode structure and schematic symbol.

Semiconductor *pn* junction structure

Metal contacts and wire leads

p   *n*

(a) Basic diode structure

Anode (A)   Cathode (K)

(b) Schematic symbol
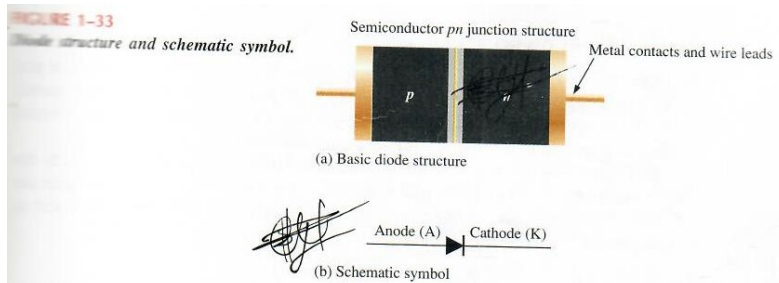
The basic DC power supply

V_{DC}

0

V_{DC}

0

V_{ac} Rectifier   Filter   Regulator   Load

FIGURE 2-1
Block diagram of a dc power supply.
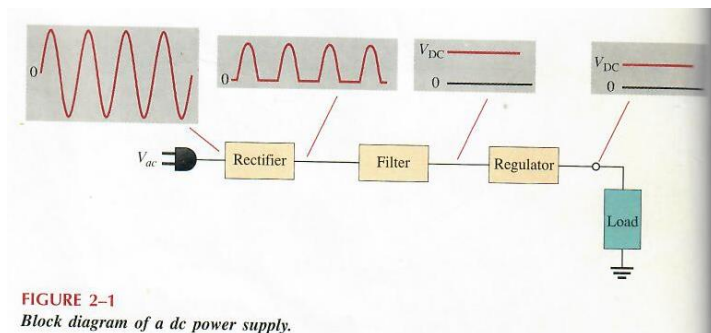
The rectifier(half wave or full wave) converts the ac input voltage to a pulsating dc voltage. The filter eliminates the fluctuations in the rectified voltage and produces a relatively smooth dc voltage. The regulator is a circuit that maintains a constant dc voltage for variations in the input line voltage or in the load.

THE BIPOLAR JUNCTION TRANSISTOR(BJT)

C (collector)

n   Base-Collector junction

B
(base)   *p*

n   Base-Emitter junction

E (emitter)

(b) *npn*

Basic
transistor
construction
Fig 4-1

C
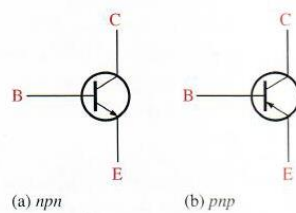
p

B   *n*

p

E

(c) *pnp*

The bipolar junction transistor is constructed with three doped semiconductor regions separated by two pn junctions. The three regions are called emitter, base and collector. There are two types of BJTs: one type consists of two n regions separated by a p region(npn)

13

and the other consists of two p regions separated by an n region(pnp). The pn junction joining the base region and the emitter region is called the base-emitter junction.

The pn junction joining the base region and the collector region is called the base-collector junction. A wire lead connects to each of the three regions(Fig 4-1). The leads are labelled E,B and C for emitter, base, and collector, respectively.

The base region is lightly doped and very thin compared to the heavily doped emitter and the moderately doped collector regions. The term bipolar refers to the use of both holes and electrons as carriers in the transistor structure.

FIGURE 4–2
Standard bipolar junction transistor (BJT) symbols.

C

B

E

(a) *npn*

C

B

E

(b) *pnp*

BASIC TRANSISTOR ACTION(NPN)

BE junction
depletion region

BC junction
depletion region

Emitter (*n*)   Base (*p*)   Collector (*n*)

Reverse-biased
BE junction

Reverse-biased
BC junction

Base electron current

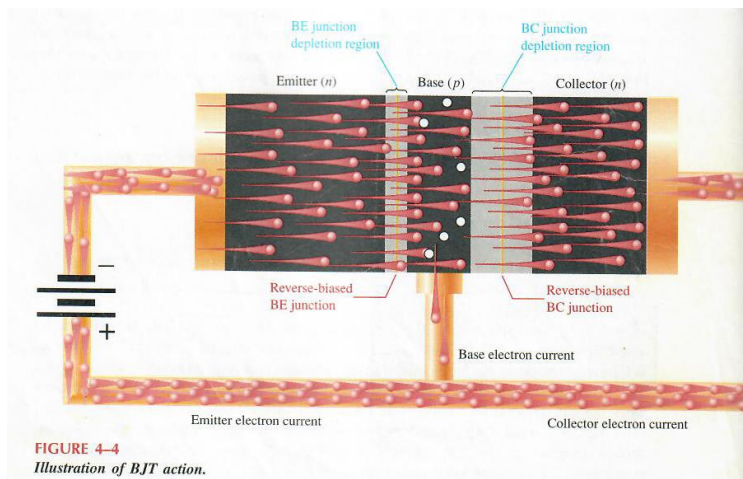Emitter electron current

Collector electron current

FIGURE 4–4
Illustration of BJT action.

Fig 4-4 shows the proper bias arrangement for the npn transistor for active operation as an amplifier. The base-emitter (BE) junction is forward-biased and the base-collector(BC) junction is reverse-biased. The forward bias from the base to emitter narrows the BE depletion region, and the reverse bias from base to collector widens the BC depletion region. The heavily doped n-type emitter is teeming with conduction-band (free) electrons that easily diffuse through the forward-biased BE junction into the p-type base region where they become minority carriers. The base region is lightly doped and very thin so that it has a very limited number of holes. Thus only a small percentage of all the electrons flowing through the BE junction can combine with the available holes in the base. These relatively

few recombined electrons flow out of the base lead as valence electrons, forming the small base electron current.

Most of the electrons flowing from the emitter into the thin, lightly doped base region do not recombine but diffuse into the BC depletion region where they are pulled through the reverse-biased BC junction by the collector supply voltage. The electrons now move through the collector lead, and into the positive terminal of the collector voltage source as the collector current.

TRANSISTOR CHARACTERISTICS AND PARAMETERS

(a) *npn*

FIGURE 4–6
Transistor dc bias circuits.
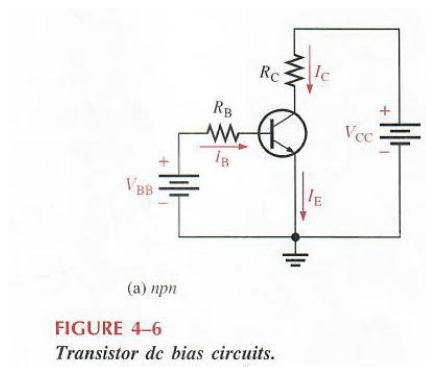
Fig 4-6 shows that the emitter current(I_E) is the sum of the collector current(I_C) and the base current(I_B) ie $I_E = I_C + I_B$ ....[1]

V_{BB} forward biases the base-emitter junction and V_{CC} reverse biases the base-collector junction.

DC Beta (β_{DC}) and DC Alpha (α_{DC})

The ratio of the collector current I_C to the base current I_B is the dc **beta** (β_{DC}), which the dc current **gain** of a transistor.

$$\beta_{DC} = \frac{I_C}{I_B} \tag{4--}$$

Typical values of β_{DC} range from less than 20 to 200 or higher. β_{DC} is usually designat as h_{FE} on transistor data sheets.

$$h_{FE} = \beta_{DC}$$

The ratio of the collector current I_C to the emitter current I_E is the dc **alpha** (α_{DC}).

$$\alpha_{DC} = \frac{I_C}{I_E} \tag{4--}$$

Typically, values of α_{DC} range from 0.95 to 0.99 or greater, but α_{DC} is always less than The reason is that I_C is always slightly less than I_E by the amount of I_B. For example $I_E = 100$ mA and $I_B = 1$ mA, then $I_C = 99$ mA and $\alpha_{DC} = 0.99$.

Relationship of β_{DC} and α_{DC}

Let's start with the current formula $I_E = I_C + I_B$ and divide each current by I_C:

$$\frac{I_E}{I_C} = \frac{I_C}{I_C} + \frac{I_B}{I_C} = 1 + \frac{I_B}{I_C}$$

Since $\beta_{DC} = I_C/I_B$ and $\alpha_{DC} = I_C/I_E$, we can substitute the reciprocals into the equation:

$$\frac{1}{\alpha_{DC}} = 1 + \frac{1}{\beta_{DC}}$$

By rearranging and solving for β_{DC}, we get

$$\frac{1}{\alpha_{DC}} = \frac{\beta_{DC} + 1}{\beta_{DC}}$$
$$\beta_{DC} = \alpha_{DC}(\beta_{DC} + 1)$$
$$\beta_{DC} = \alpha_{DC}\beta_{DC} + \alpha_{DC}$$
$$\beta_{DC} - \alpha_{DC}\beta_{DC} = \alpha_{DC}$$
$$\beta_{DC}(1 - \alpha_{DC}) = \alpha_{DC}$$

$$\beta_{DC} = \frac{\alpha_{DC}}{1 - \alpha_{DC}} \tag{4-}$$

Equation (4--4) shows that the closer α_{DC} is to 1, the higher the value of β_{DC}.

16

Current and Voltage Analysis

Consider the circuit configuration in Figure 4–7. Three transistor currents voltages can be identified.

I_B: base current (dc)

I_E: emitter current (dc)

I_C: collector current (dc)

V_{BE}: dc voltage at base with respect to emitter

V_{CB}: dc voltage at collector with respect to base

V_{CE}: dc voltage at collector with respect to emitter

FIGURE 4–7
Transistor currents and voltages.
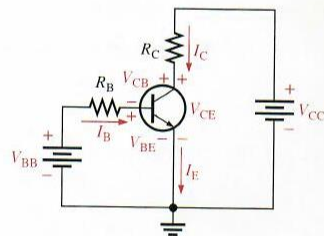
V_{BB} forward-biases the base-emitter junction and V_{CC} reverse-bias collector junction. When the base-emitter junction is forward-biased, it is li biased diode and has a forward voltage drop of

$$V_{BE} \cong 0.7 \text{ V}$$

Since the emitter is at ground (0 V), by Kirchhoff's voltage law, the voltage a

$$V_{R_B} = V_{BB} - V_{BE}$$

Also, by Ohm's law,

$$V_{R_B} = I_B R_B$$

Substituting for V_{R_B} yields

$$I_B R_B = V_{BB} - V_{BE}$$

Solving for I_B, you get

$$I_B = \frac{V_{BB} - V_{BE}}{R_B}$$ (
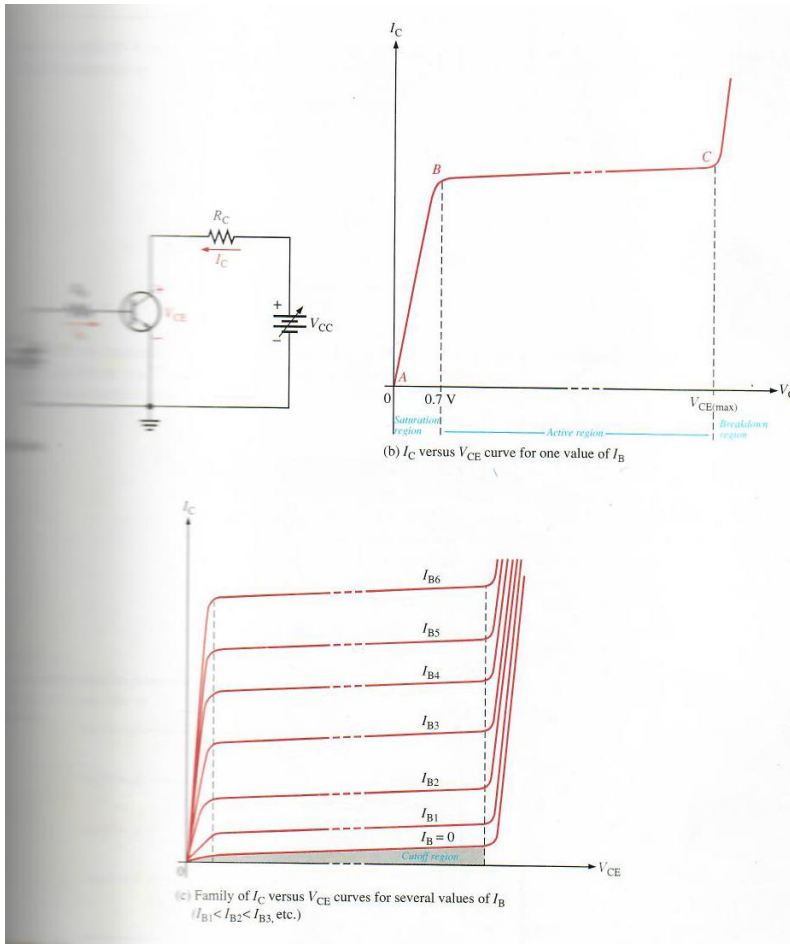
The drop across R_C is

$$V_{R_C} = I_C R_C$$

The voltage at the collector with respect to the emitter, which is grounded, is

$$V_{CE} = V_{CC} - I_C R_C$$ (

where $I_C = \beta_{DC} I_B$. The voltage across the reverse-biased collector-base junction is

$$V_{CB} = V_{CE} - V_{BE}$$ (

18

Collector characteristic curves

(b) I_C versus V_{CE} curve for one value of I_B

(c) Family of I_C versus V_{CE} curves for several values of I_B
($I_{B1} < I_{B2} < I_{B3}$, etc.)

Using circuit in Fig 4-9a a set of collector characteristic curves are generated that show how the collector current, I_C, varies with the collector-to-emitter voltage, V_{CE}, for specified values of base current, I_B.(Fig 4-9c) V_{BB} and V_{CC} are variable sources of voltage.

There are three distinct regions:

Saturation region: as V_{CC} is increased ,V_{CE} increases gradually as the collector current I_C increases between A and B.

Linear region: I_C levels off and remains essentially constant for a given value of I_B as V_{CE} continues to increase between B and C. in this region the value of IC is determined from $I_C = \beta_{DC}I_B$.

19

Breakdown region: the collector current increases rapidly with little increase in V_{CE}.(to the right of C)

Cutoff region: When $I_B = 0$, there is a small collector leakage current.

THERMODYNAMICS

It is the study of thermal energy (or internal energy) of systems .Temperature is one of the seven SI base quantities. Physicists measure temperature on the Kelvin scale in Kelvins.

THE ZEROTH LAW OF THERMODYNAMICS

If bodies A and B are each in thermal equilibrium with a third body T, then they are in thermal equilibrium with each other. It means every body has a temperature. When two bodies are in thermal equilibrium their temperatures are equal. We use a thermometer to measure temperature.

MEASURING TEMPERATURE

The triple point is a standard fixed point where solid ice, liquid water and water vapour coexist in equilibrium. The temperature of this mixture has been defined as 273,16 K

(T_3 =273,16 K) the subscript 3 means triple point. The Kelvin is set as $\frac{1}{273.16}$ of the difference between absolute zero and triple point temperature of water.

THE CONSTANT- VOLUME GAS THERMOMETER
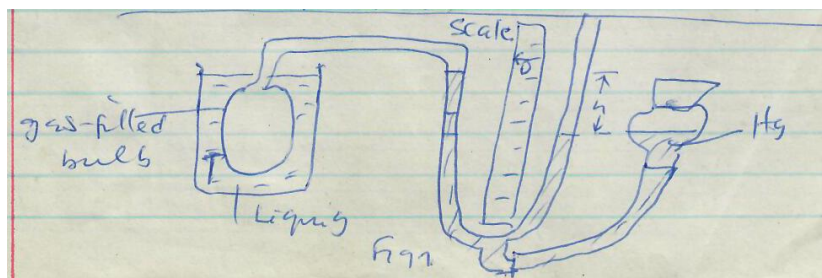
This is a standard thermometer against which all other thermometers are calibrated. The constant volume gas thermometer consists of a gas filled bulb connected by a tube to a mercury manometer. The Hg level can be brought to zero by raising or lowering the reservoir R, to maintain a constant volume. Temperature of any body in thermal contact with the bulb b is $T = Cp$..[1]

C is a constant and p is pressure of gas in bulb

$p = p_o - \rho gh$..[2]

P_o is atmospheric pressure, ρ is Hg density, h is difference of Hg levels. If bulb is next put in a triple – point cell the temperature being measured is:

$T_3 = Cp_3$..[3]

21

where p_3 is the gas pressure.

From [1] and [3] $T = T_3(\frac{p}{p_3})$=273.16 $K(\frac{p}{p_3})$...[4]

If we use different gases they will register different temperatures but as p→o the temperature converges for all gases.

The commented note:

$$T = 273.16\ K(\frac{lim}{gas\to 0}\frac{p}{p_3})\text{...[5]}$$

THE CELSIUS AND FAHREINHEIT SCALE

$T_c = T - 273.15^0$...[6]

$T_F = \frac{9}{5}T_C + 32^0$..[7]

Problem : On a temperature scale called Z, on which the boiling point is $65,0^0Z$ and the freezing point is $-14,0^0Z$.To what temperature on the Fahrenheit scale would a temperature of T =$-98,0^0Z$ correspond? Assuming Z scale is linear?

T is below freezing point by -14 -(-98)=84^0Z

The temperature difference of freezing and boiling point is 65-(-14)=79^0Z. This corresponds to 212^0F -32^0F=180^0F on the Fahreinheit scale.

84^0Z corresponds to $84^0Zx\frac{180^0}{79^0Z}{}^0F$= 191^0F

Therefore T=32^0-191^0= -159^0FzzZZZz

THERMAL EXPANSION

1 .Linear Expansion: If temperature of metal rod of length L is raised by ΔT its length L increases by   ΔL= LαΔT...[8]

α is coefficient t of linear expansion (units are $^0C^{-1}$ or K^{-1} and depends on material).

Commented [C6]: As the pressure in the bulb turn to zero

Commented [C7]: 32 is the increasing point of wateron the F scale

22

2. Volume Expansion : If volume of solid's temperature is raised by ΔT the volume is raised by ΔV = VβΔT, β is coefficient of volume expansion. β= 3α………………….[9]

Water does not behave like other liquids. Between 0^0C and 4^0C water contracts and starts expanding thereafter. Thus at 4^0C water passes through a maximum density.

TEMPERATURE AND HEAT

The temperature of a cup of hot coffee, left sitting on the table falls until it reaches room temperature.
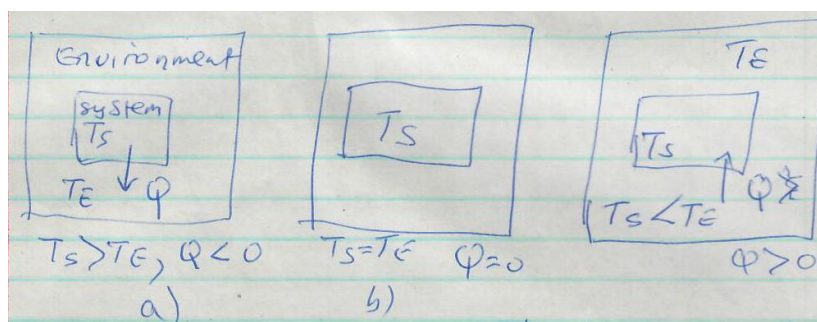
(a) Temperature changes by transfer of energy between thermal energy of the system (hot coffee) and the system's environment until thermal equilibrium is reached as in (b).

-Thermal energy is an internal energy that consists of the kinetic and potential energies as associated with the random motions of the atoms, molecules and other microscopic bodies within an object. The transferred energy is called heat, Q. Q is +ve when energy transferred to a system as thermal energy from its environment and – ve when transferred from system to its environment (heat is lost).

In (b) heat is neither absorbed nor released; Q = 0 and $T_S = T_E$.

In (c) heat is absorbed when $T_S < T_E$ and Q is +ve.

-Heat is the energy that is transferred between a system and its environment because of a temperature difference that exists between them. The unit of heat is the joule (J)

THE ABSORPTION OF HEAT BY SOLIDS AND LIQUIDS.aaaaa

Heat capacity, C is the amount of heat an object absorbs or releases per unit degree Celsius. Units are J^0C^{-1} or JK^{-1}.

Q = CΔT = C(T_f -T_i)……………………………………………………………………………….[10]

T_f is final temperature and T_i is initial temperature .

Specific heat capacity, c is heat an object absorbs /releases per unit mass per unit degree C or Kelvin. $Q = mc\Delta T = mc(T_f - T_i)$..[11] Units are $Jkg^{-1}C^{-1}$ or $Jkg^{-1}K^{-1}$.

Molar specific heat –Heat an object absorbs or releases per unit mole per unit degree Celsius /Kelvin .

1 mole = $6,02x10^{23}$ elementary units. e.g 1 mole of Al contains $6,02x10^{23}$ atoms and 1 mole of Al_2O_3 contains $6,02 x10^{23}$ molecules.

-In determining and using specific heat of any substance we need to know the conditions under which the energy is transferred eg at constant volume or pressure.

HEATS OF TRANSFORMATION

When energy is absorbed by a solid or a liquid the temperature of the sample does not necessarily rise. The sample may change from one phase into another. Matter exists in three states:

Solid state –molecules of a sample are locked into a fairly rigid structure by their mutual attraction.

Liquid state – molecules have more energy and move about more, they don't have regular structure can flow or settle in a container.

Gas or vapour state –molecules have even more energy, are free of one another and can fill up the full volume of the container.

To melt a solid means to change it from solid to liquid state. The process requires energy e.g melting an ice cube.

-To vaporise a liquid means to change it from liquid state to the vapour or gas state – energy is required.

-Condensing a gas to form a liquid- energy is removed from the gas.

HEAT OF TRANSFORMATION (L) – the amount of energy per unit mass that must be transferred as heat when a sample under goes a phase change completely.

$Q = Lm$..[12]

Heat of fusion L_f – phase change is from solid to liquid (sample must absorb heat)

Heat of vaporisation L_v – phase change from liquid to gas (sample must absorb heat).

-When phase change is reversed heat is released.Units are in JKg^{-1}

Example:

24

How much heat must be absorbed by ice of mass m =720 g at -10 ^{0}C to take it to liquid state at 15^{0}C.

Total heat required $Q_{tot} = Q_1+Q_2+Q_3$

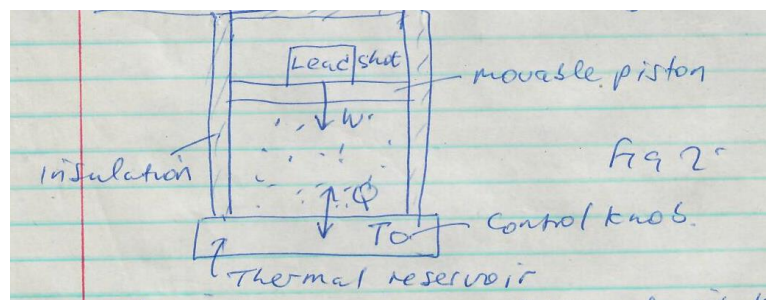Q_1 =c_{ice}m $(T_f -T_i)$ heat required to raise the temperature of ice from -10 C to 0 C

Q_2 =L_fm heat required to melt the ice.

Q_3= c_{liq}m$(T_f -T_i)$ heat required to raise the temperature of the water from 0^0 C to15^{0}C

Q=(2220)(0,720Kg)(0 –(-10)+(333000)(0,720)+(4190) (0,720 (15 -0)

=45252J=45,3KJ

HEAT AND WORK (energy transfer)

Consider a system – a gas confined to a cylinder with a moveable piston (Fig 2).The upward force on the piston due to the pressure of the confined gas is equal to weight of lead shot loaded on top of the piston. The system is insulated from surroundings- doesn't allow transfer of energy as heat. The bottom of cylinder rests on a thermal reservoir (eg hot plate) whose temperature can be controlled by a knob T.

The system starts from an initial state i, described by pressure P_i, volume V_i and a temperature T_i and changes to a final state described by P_f, V_f and T_f. The procedure by which you change the system from its initial state is called a thermodynamic process.

-Energy may be transferred into a system from thermal reservoir +ve heat or vice versa –ve heat.

-Work can be done by system to raise loaded piston +ve work or lower it –ve work.

-We assume the change is done so slowly so that every part of the system is in thermal equilibrium with one another. Suppose you remove a few lead shot allowing the gas to push upwards the piston through a differential displacement ds with an upward force F

F = pA..[13](A is area of piston and p is pressure of gas)

Differential work dW done by the gas during the displacement is :

dW =F. d s = (Pa) (ds)= pAds

= pdV..[14]

Where dV is differential change of the volume of the gas owing to movement of piston.

Total work done by the gas is: W = = $\int dW = \int_{Vi}^{Vf} pdV$ ....[15]

During the change in volume, pressure and temperature of the gas may change. There are several ways to take the gas from i to f.This can be illustrated in a P – V diagram.

(a) Pressure decreases as volume increases.
Work done is represented by the shaded area under the curve between points i and f.
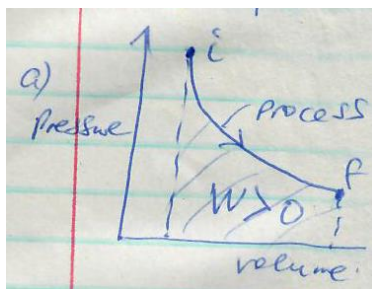
The work done is + ve.
(b) Two way process: ia at constant pressure, af at constant volume. You raise volume from V_i to V_f by slowly raising temperature to some higher value T_a. +ve work is done by expanding gas to lift piston upward. Heat is +ve because it is added to system in this process. Then you use control knob to lower temperature and pressure drops from P_a to P_f. Here heat is lost to reservoir by system. For iaf +ve work W is carried out only during step ia(represented by shaded area). Step af is done at constant volume(so you must wedge piston) Energy is transferred as heat for ia and af with a net energy transfer Q.

26
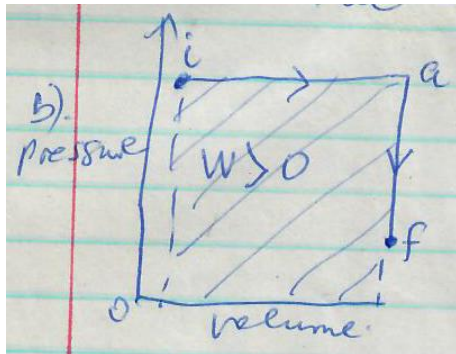
(c) The processes in (b) are done in reverse order. W is now smaller than in (b) as is the net heat absorbed.
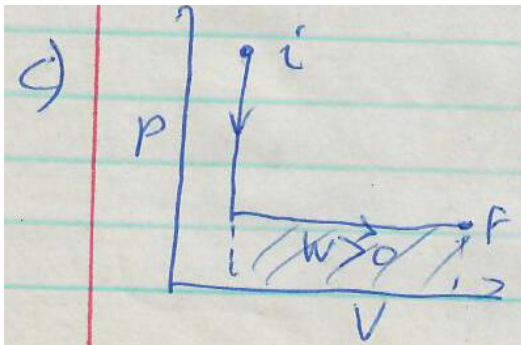
(d) You can make the work done by gas as small as you want (icdf) or as large as possible(ighf).
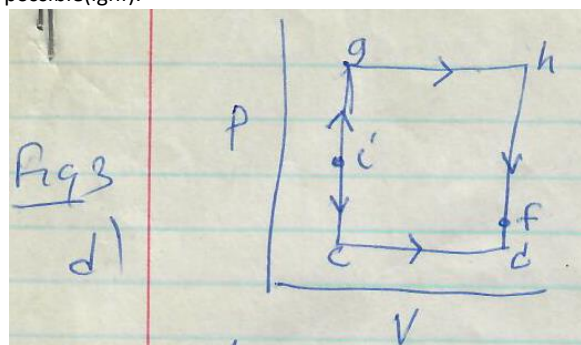
A system can be taken from a given initial state to a given final state by an inifinite number of processes. Heat may or may not be involved and generally W and Q will have different values for different processes. Heat and work are path independent quantities.

(e) -ve work is done by a system as some external force compresses the system, reducing volume.
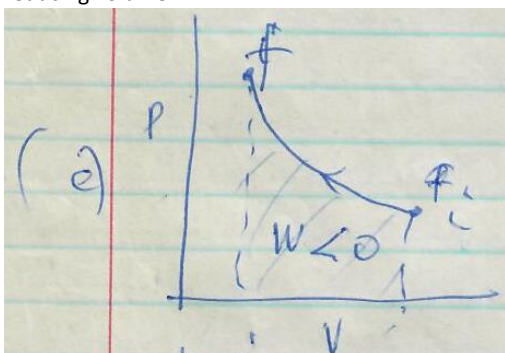
(f) A thermodynamic cycle in which the system is taken from some initial state i to some final state f and then back to i.
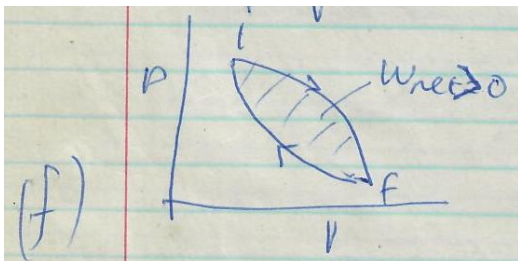
Note work is   + ve work during expansion and – ve work during compression.

THE FIRST LAW OF THERMODYNAMICS

The quantity Q –W is the same for all processes but Q and W depend on the nature of the process. Q –W only depend on the initial and final states and not how we get there. Combinations of Q and W, Q alone, W alone, Q+W and Q -2 W are path dependent but not Q-W.

Change in internal energy : $\Delta E_{int} = E_{intf} - E_{inti}$=Q-W   (1st law)..................[16]

If the thermodynamic system undergoes a differential change we can write:

$$dE_{int} = dQ - dW \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\text{[17] 1}^{st}\text{ law.}$$

The internal energy of a system E_{int} tends to increase if energy is added as heat Q and tends to decrease if energy is lost as work W by the system.

-The first law of thermodynamics is an extension of the principle of conservation of energy when the system is not isolated. In the first law we assume ΔK = ΔU = 0 where ΔK is kinetic energy and ΔU is potential energy.

The work done on a system is always the –ve of the work done by the system so we can write: $\Delta E_{int} = Q + W_{on}$(work done on system)

SOME SPECIAL CASES OF THE FIRST LAW OF THERMODYNAMICS.

1. Adiabatic process - is one which occurs where no transfer of energy of heat occurs ie Q = 0. ΔE_{int} = -W …………………………………………………………………………………[18] adiabatic process

This means if   work is done by the system (W is +ve), the internal energy decreases by the amount of work and if work is done on system internal energy increases by that amount.

2.Constant – Volume process. –If the volume of a system is held constant, the system can do no work, ie W = 0, ΔE_{int} = Q……………………………………………………………..[19] (constant – volume process). ie if Q is absorbed by system, ΔE_{int} increases and vice versa.

3. Cyclical process - After certain interchanges system is restored to its initial state.

   ΔE_{int} = 0. Therefore Q = W……………………………………………………………[20] cyclic process.

4. Free expansion Q = W = 0 , ΔE_{int} = 0……………………………………………….[21]

HEAT TRANSFER MECHANISMS.
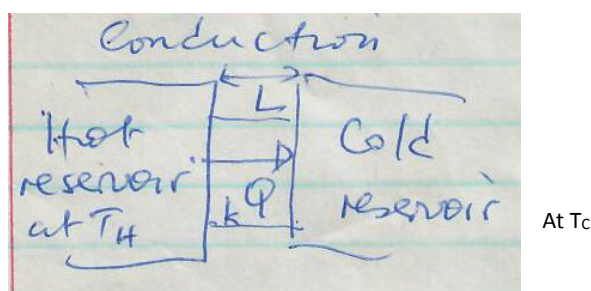
Consider a slab of face area A thickness L whose faces are maintained at temperatures T_H

and T_C by a hot and cold reservoir respectively. If Q is energy transferred as heat through the slab from its hot face to its cold face in time t

Conduction rate $P_{cond} = \dfrac{Q}{t} = kA\dfrac{T_H - T_C}{L}$..[22]

k,the thermal conductivity is a constant which depends on the material of which the slab is made. Thermal resistance to conduction (R – value)

$R = \dfrac{L}{k}$...[23]

A high R value makes a material a poor thermal conductor and therefore a good thermal insulator.
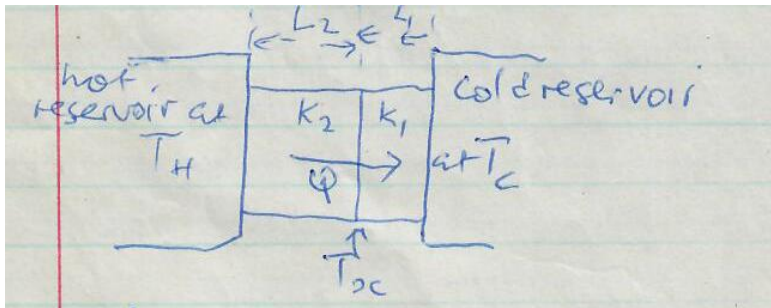
In the steady state, the conduction rates through the 2 materials must be equal. Let T_X be the interface temperature.

$P_{cond} = k_2 A\dfrac{(T_H - T_X)}{L_2} = k_1 A\dfrac{(T_X - T_C)}{L_1}$...[24]

$T_X = \dfrac{k_1 L_2 T_C + k_2 L_1 T_H}{k_1 L_2 + k_2 L_1}$..[25]

$P_{cond} = \dfrac{A(T_H - T_C)}{\dfrac{L_1}{k_1} + \dfrac{L_2}{k_2}}$...[26]

Radiation – Energy is transferred via electromagnetic waves as thermal radiation. The rate at which a body radiates energy is given by $P_{rad} = \sigma\varepsilon A T^4$ ....[27]

σ-Stefan – Boltzmann constant

A-Surface area of object

T– temperature of object

ε - emissivity of the object's surface

Net rate of exchange of heat with the environment is given by ;

$P_{net} = P_{abs} - P_{rad} = \sigma\varepsilon A(T_{env}^4 - T^4)$...[28]

Where T_{env} is temperature of environment

KINETIC THEORY OF GASES

A gas is made up of moving atoms or molecules. The pressure exerted by a gas is related to the collisions of its molecules with the walls of its container. A gas is able to fill the volume of its container due to the freedom of motion of its molecules, and the temperature and internal energy of a gas must be related to the kinetic energy of these molecules.

Avogadro's Number, N_A

One mole is the number of atoms in a 12 g sample of carbon 12. $N_A = 6{,}02 \times 10^{23}$ mol^{-1}[1]

Number of moles n in a sample containing N molecules is $n = \dfrac{N}{N_A}$..............................[2]

We can also express number of moles in a sample as $n = \dfrac{M_{sum}}{M}$....................................[3]

$$= \frac{M_{sum}}{mN_A}$$ ....[4]

where M_{sum} is mass of sample, M is molar mass and m is mass of 1 molecule.

Ideal gas

At low enough densities, all real gases tend to obey the relation.

$\qquad$ PV = nRT(ideal gas law)..[5]

R is the gas constant, R =8,31J/mol.K..[6]

T is temperature in Kelvin .

Equation (5) can be rewritten in another form:

Boltzmann constant $k = \dfrac{R}{N_A}$,$\rightarrow R = kN_A$..[7]

nR = knN_A= kN and PV = NkT...[8]

k=8.31/6.02x10^{23}=1.38x10^{-23}JK^{-1}

Work done by an ideal gas

Suppose we allow a gas to expand from an initial volume V_i to a final volume V_f while we keep the temperature of the gas constant. The process at constant temperature is called an isothermal expansion and (reverse is an isothermal compression). On a PV diagram an isotherm is a curve that connects points that have the same temperature.

Fig 4 — Three isotherms on a PV diagram

(figure handwritten notes: "Isotherm", "have the same temperature", "T=320k", "f T=310k", "T=300k")
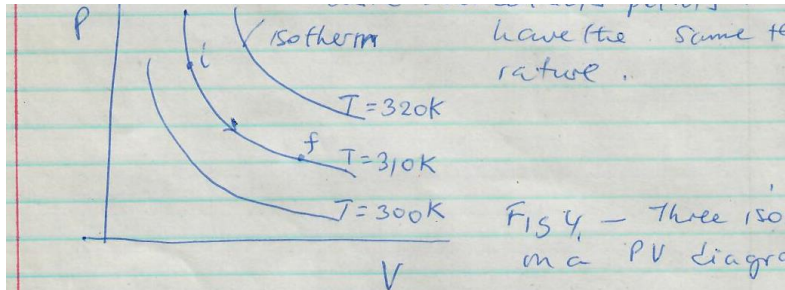
From [1] $P = nRT\frac{1}{V}=(a\ constant)\frac{1}{V}$ ....[9]

To find the work done by an ideal gas during an isothermal expansion:

$W = \int_{V_i}^{V_f} pdV$ ....[10]

$= \int_{V_i}^{V_f} \frac{nRT}{V}dV$

$= nRT \int_{V_i}^{V_f} \frac{dV}{V} = nRT ln\frac{V_f}{V_i}$ ....[11]

Work done in an isothermal expansion is +ve while that done on a compressional expansion is − ve.

Work done at constant volume and constant pressure:

For a constant-volume process, volume of gas is constant then [10] yields

$$W = p \int_{V_i}^{V_f} dV = 0 \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots [12]$$

For a constant pressure process

$W = \int_{V_i}^{V_f} pdV = p(V_f − V_i)=p\Delta V$ ....[13]

Pressure, temperature and RMS speed

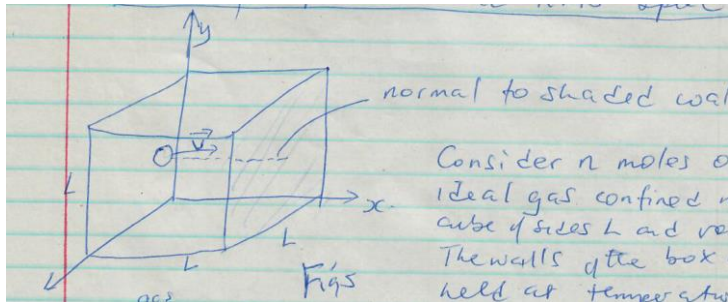Consider n moles of an ideal gas confined in a cube of sides L and volume V. The walls of the box are held at temperature T. A gas molecule mass m and velocity v is about to collide with shaded wall elastically. Considering the x component velocity v_x

Change in momentum $\Delta p_x = (-mv_x - mv_x) = -2mv_x$ ....[14]

Momentum delivered to wall = +2mv$_x$

Time between collisions $\Delta t = \dfrac{2L}{v_x}$

Rated change of momentum $= \dfrac{\Delta p_x}{\Delta t} = \dfrac{2mv_x}{2L/v_x}$ But $F = \dfrac{dp}{dt}$, $\dfrac{\Delta p_x}{\Delta t} = \dfrac{2mv_x}{2L/v_x} = \dfrac{mv_x^2}{L}$

Pressure on wall $p = \dfrac{F_x}{L^2} = (m\dfrac{v_{x1}^2}{L} + m\dfrac{v_{x2}^2}{L} + \ldots \ldots + m\dfrac{v_{xN}^2}{L})/L^2$ ....[15]

$= (\dfrac{m}{L^3})(v_{x1}^2 + v_{x2}^2 + \cdots \ldots \ldots + v_{xN}^2)$ ....[16]

Where N is number of molecules in box $N = nN_A$

Pressure: $p = \dfrac{nmN_A}{L^3}(v_x^2)_{avg}$ ....[17]

For any molecule $v^2 = v_x^2 + v_y^2 + v_z^2$

Assuming all components are equal $v_x = v_y = v_z$

$$v_x^2 = \frac{1}{3}v^2$$

$$p = \frac{nM(v^2)_{avg}}{3V}$$

$$pV = \frac{nM(v^2)_{avg}}{3} = nRT$$

$$v_{rms} = \sqrt{\frac{3RT}{M}}$$ ....[18]

Translational Kinetic Energy

The average translational kinetic energy of a molecule

$$K_{avg} = (\tfrac{1}{2}mv^2)_{avg} = \tfrac{1}{2}mv_{rms}^2$$

$$= (\tfrac{1}{2}m)(\tfrac{3RT}{M}).$$

$$= \frac{3RT}{2N_A} \qquad (\tfrac{M}{m} = N_A)$$

$$= \frac{3}{2}kT$$ ....[19]

$$(k = \frac{R}{N_A})$$

At a given temperature T, all ideal gas molecules no matter what their mass – have the same average translational kinetic energy namely $\frac{3}{2}kT$

Mean Free Path, λ

Molecules move in a zig- zag manner but in between collisions with one another they travel in straight lines at constant speed. The mean free path, λ is the average distance travelled by a molecule between collisions.

-We expect λ to vary inversely as the density of the molecules or N/V.

-We also expect λ to vary inversely with the size of the molecules, with their diameter.

-We can also predict that λ should vary inversely on square of molecular diameter.

$$\lambda = \frac{1}{\sqrt{2}\pi d^2 N/V}$$ ....[20] (mean free path)

The distribution of molecular speeds

The root mean square speed v_{rms} gives us a general idea of molecular speeds in a gas at a given temperature.

The probability of a speed at a certain temperature is given by Maxwell's speed distribution law.

$$P(v) = 4\pi(\frac{M}{2\pi RT})^{3/2}v^2 e^{-Mv^2/2RT}$$ ....[21]

$$P(v) = 4\pi\left(\frac{M}{2\pi RT}\right)^{3/2} v^2 e^{-Mv^2/2RT} \qquad (21)$$

Maxwell's speed distribution for O molecules at 300 K

Fig 6

Fig 7 Curves for 300k and 80k
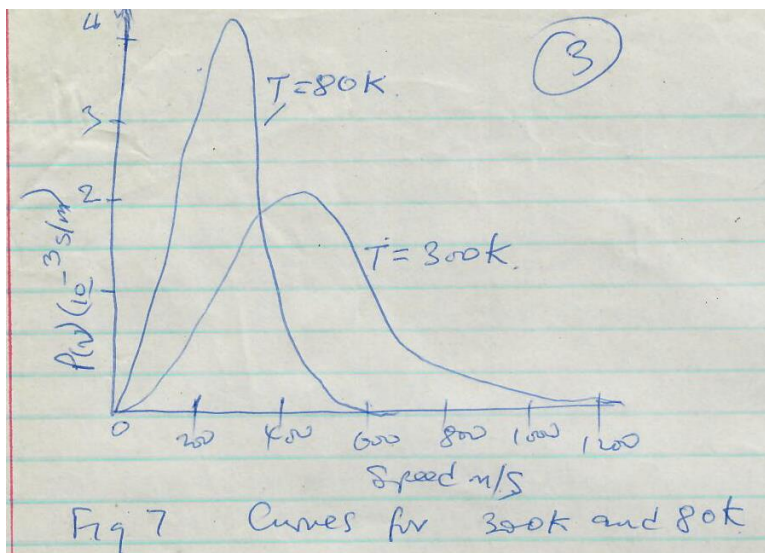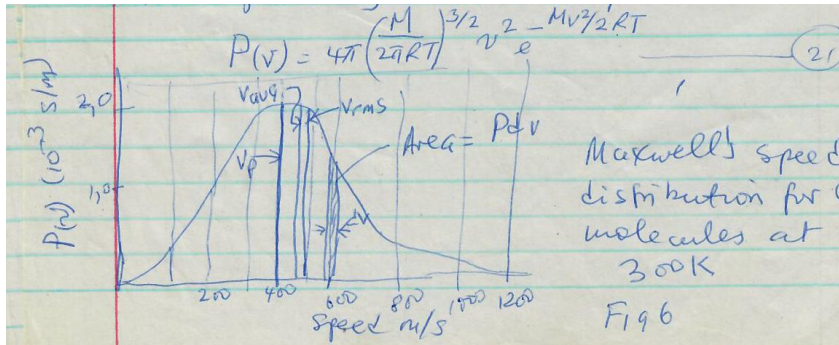
P(v) is a probability distribution function

P(v)dv (dimensionless quantity) is the fraction of the molecules whose speeds lie in the interval dv centred on speed v.

$\int_0^\infty P(v)dv = 1$ is total area under distribution curve.

Fraction of molecules with speeds in interval v_1 to v_2 is then

Frac $= \int_{v_1}^{v_2} P(v)dv$

Average, RMS and most probable speeds.

Average speed, $v_{avg} = \int_0^\infty vP(v)dv = \sqrt{\dfrac{8RT}{\pi M}}$ ....[22]

35

$$(v^2)_{avg} = \int_0^\infty v^2 P(v)dv = \frac{3RT}{M}$$

$$\sqrt{(v^2)_{avg}} = v_{rms} = \sqrt{\frac{3RT}{M}}$$...[23]

Most probable speed v_P is speed at which P is maximum. To calculate v_P, we set $\frac{dP}{dv} = 0$

$$v_P = \sqrt{\frac{2RT}{M}}$$...[24]

While most molecules have speeds v_P, there are some with even higher speeds in the tail.

The Molar specific heats of an ideal gas

Assume ideal gas is monatomic eg He, Ne or Ar and E_{int} = sum of the translational kinetic energies of its atoms, $K_{avg} = \frac{3}{2}kT$

A sample of n moles of such a gas has nN_A atoms. The internal energy of sample:

$E_{int} = nN_A K_{avg} = nN_A(\frac{3}{2}kT)$...[25]

$= \frac{3}{2}nRT$...................................(monatomic ideal gas)...[26]
(But $k=R/N_A$)

The internal energy E_{int} of an ideal gas is a function of the gas temperature only; it does not depend on any other variable.
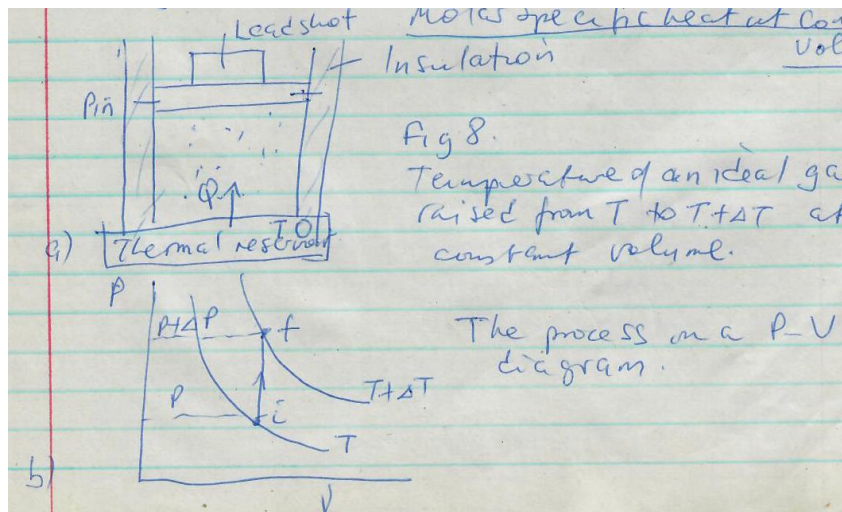
Molar specific heat at constant volume

Fig 8a shows n moles of an ideal gas at pressure P, temperature T, confined to a cylinder of volume V.

If a small amount of heat is added temperature rises by a small amount to T + ΔT and pressure rises to P + ΔP bringing gas to final state f.

$$Q = nC_v \Delta T \text{ (constant volume)} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[27]$$

C_v- molar specific heat at constant volume substituting into first law thermodynamics.

$\Delta E_{int} = Q - W \ yields$

$\Delta E_{int} = nC_v \Delta T - W$ ....[28]

$and \ C_v = \dfrac{\Delta E_{int}}{n \Delta T}$...[29]

$But \ \Delta E_{int} = \dfrac{3}{2} nR \Delta T \ \ Therefore \ C_v = \dfrac{3}{2} R$=12.55 J/mol.K................[30]

This kinetic theory prediction for ideal gases experimentally agrees very well for monatomic gases. For any ideal gas $E_{int} = \dfrac{3}{2} nRT = nC_v T$...[31]

Therefore $\Delta E_{int} = nC_v \Delta T$(ideal gas, any process)

For diatomic gases (ideal gas) $C_v = \dfrac{5}{2} R$ =20.8 J/mol.K

For polyatomic gases (ideal gas) $C_v = 3R = 24.9$ J/mol.K

*A change in the internal energy of a confined gas E_{int} depends on the change in the gas temperature only; it does not depend on the type of process that produces the temperature change as shown in Fig 9.

Fig 9.

Path 1 — ⊆ constant - volume process
Path 2 —   "   - pressure   "
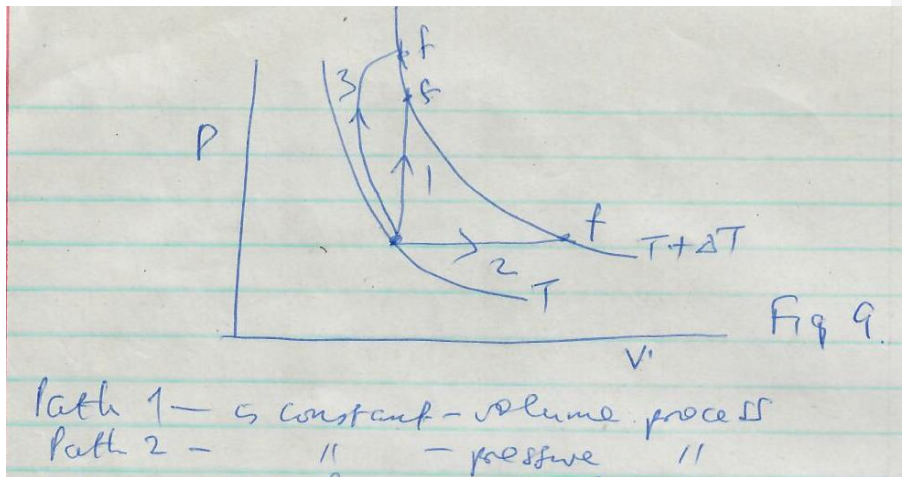
Although the values of Q (heat) and work W associated paths differ as V_f and P_f, ΔE_{int} for the 3 paths is the same since they both involve temperature change ΔT.

Molar Specific Heat at Constant Pressure

a)

b)
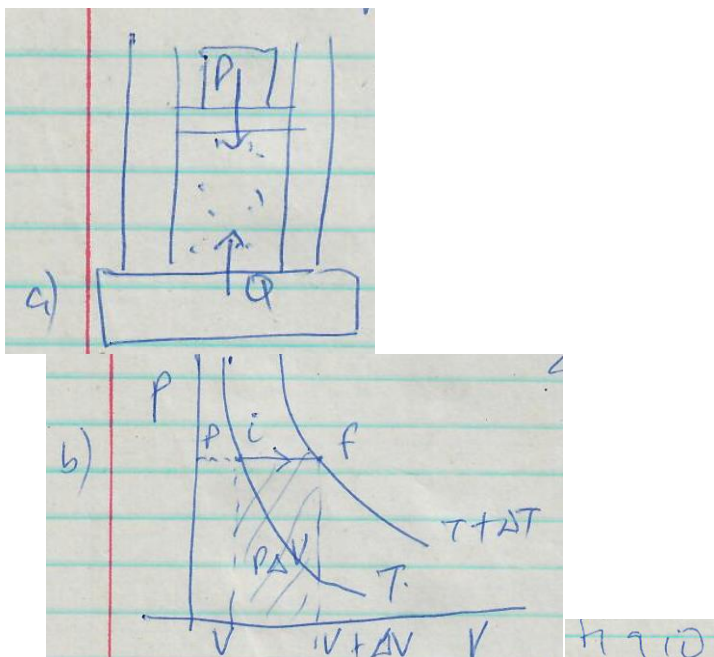
Temperature of ideal gas is raised by ΔT, under under constant pressure

$Q = nC_P\Delta T$...(constant pressure...................................[32]

and $W = p\Delta V = nR\Delta T$ Substituting in first law $\Delta E_{int} = Q - W$ yields

$\Delta E_{int} = nC_P\Delta T - nR\Delta T = nC_v\Delta T$

Therefore $C_P - C_v = R$...[33]

These predictions agree with experimental values for real gases as long as densities are sufficiently low.

The adiabatic expansion of an ideal gas

a) Volume of gas is increased by removing mass from piston. The

b) The process proceeds from i to f along an adiabat on a P-V diagram
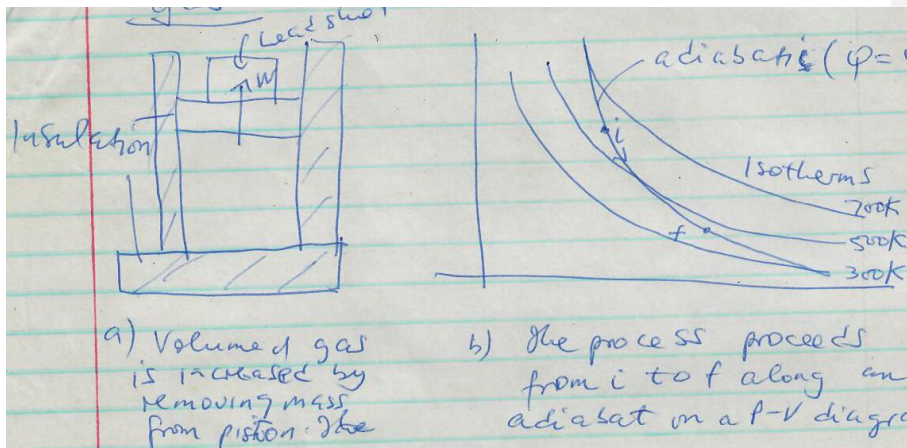
As the gas expands adiabatically the pressure and the temperature drop.

$pV^\gamma = a$ constant (adiabatic process)..[34]

Where $\gamma = \dfrac{C_P}{C_v}$

$P_i V_i^\gamma = P_f V_f^\gamma$ (PV=nRT)

$\left(\dfrac{nRT}{V}\right)V^\gamma =$ a constant

$TV^\gamma =$ a constant (adiabatic process)..[35]

$$T_i V_i^\gamma = T_f V_f^\gamma$$

Proof of (34)

If you remove some shot from piston, it will push upward and volume increases by dV since volume change is small pressure p of gas on piston is constant during the change.

$dE_{int} = Q - PdV$...[36]

39

$C_v ndT = -PdV$ and $ndT = -(\frac{P}{C_v})dV$ ....[37]

From ideal gas law $PV = nRT$ we have

$PdV + Vdp = nRdT = n(C_P - C_V)dT$

$ndT = \frac{PdV + VdP}{C_P - C_V}$ ....[38]

Equating (37) and (38)

$$\frac{dp}{p} + \left(\frac{C_P}{C_V}\right)\frac{dV}{V} = 0$$

$$lnp + \frac{C_P}{C_V}lnV = a\ constant$$

$$lnp + \gamma lnV = aconstant$$

Therefore PV^γ = a constant...[39]

Free expansion

$\Delta E_{int} = 0$

$T_i = T_f$ ....free expansion...................................[40]

$P_i V_i = P_f V_f$ ....free expansion...................................[41]

ENERGY

Kinetic energy is energy associated with stated motion of an object.

$K = \frac{1}{2}mv^2$ ....[1]

Units Joule(J).

Work, W is energy transferred to or from an object by means of a force acting on the object. Energy transferred to the object is + ve work and energy transferred from the object is – ve work. Units of work - Joule(J).

Work and Kinetic energy

40
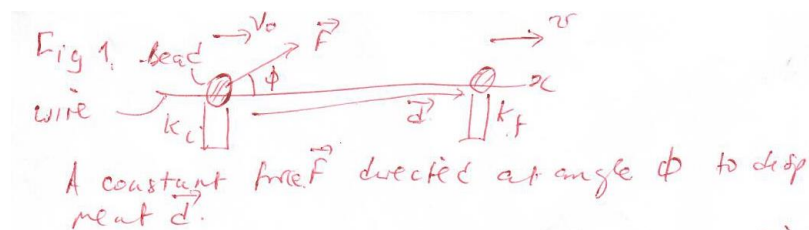
A   constant force F directed at angle Φ to displacement d. Consider a bead that can slide along a frictionless wire, stretched along a horizontal x axis. A constant force F accelerates bead mass m along the wire.

Using Newton's second law for component, along x axis.

$$F_x = ma_x \qquad \text{..}[2]$$

The force changes bead's velocity from v_o to v.

$$v^2 = v_o^2 + 2a_x d \text{..}[3]$$

$$[3]x\frac{1}{2}m \qquad \frac{1}{2}m \ v^2 = \frac{1}{2}m \ v_o^2 + F_x d \text{...}[4]$$

$$Final \ KE, K_f = \frac{1}{2}mv^2.$$

$$Initial \ KE, K_i = \frac{1}{2}mv_o^2.$$

Work done W on the bead by the force (energy transfer due to the force) is

$$W = F_x d \text{...}[5]$$

$$W = Fdcos\Phi. \ \text{(work done by a constant force from FIG 1......................}[6]$$

$$W = \vec{F}.\vec{d} \text{...}[7]$$

There are 2 restrictions to the use of [7]:

1.The force must be a constant force.

2. The object must be particle –like, it must be rigid.

A force does + ve work when it has a vector component in the same direction as the displacement and it does – ve work when it has a vector component in the opposite direction. It does zero work when it has no such vector component.

Work – Kinetic Energy Theorem

For FIG 1 change in K.E, $\Delta K = K_f - K_i = W$ ....[8] which says (change in the KE of a particle)=(network done on the particle).

We can also write $K_f = K_i + W$ ....[9]

(KE after the net work is done)= (KE before the net work)+(the net work done)

These statements are known as the work – kinetic energy theorem.

Net work W done on object is the sum of works done by individual forces or find work done by the net force.

Work done by a gravitational force

$$W_g = mgdcos\Phi \text{..}[10]$$

Work done in lifting and lowering an object

If we lift an object our applied force does work W_a on object while gravitational force tends to transfer energy from it (W_g).

The change in KE is $\Delta K = K_f - K_i = Wa + Wg$..[11]

If you lift a book from the floor to a shelf K_f and K_i are both zero and (11) reduces to :

$Wa + Wg = 0$ and $Wa = -Wg$..[12]

Work done by a spring

The spring force $F = -kx$........................Hooke's law................................[13] k is spring constant.

Work done by a spring is

$$W_s = \int_{xi}^{xf} Fdx$$...[14]

x_i and x_f are initial and final displacements.

$$W_s = \int_{xi}^{xf}(-kx)dx = -k\int_{xi}^{xf} xdx = (-\frac{1}{2}k)(x_f^2 - x_i^2)$$..................[15]

Work done by a variable Force

Fig 2
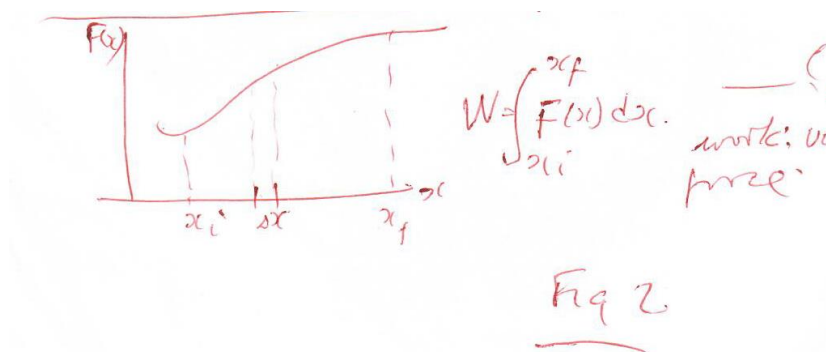
Collisions

A collision is an isolated event in which 2 or more bodies exert relatively strong forces on each other for a relatively short time.
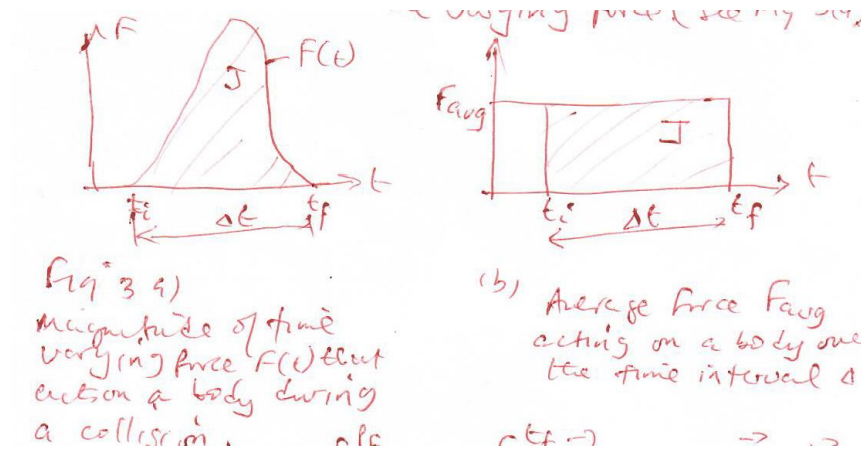
Impulse and Linear Momentum

Single collision

A force pair F(t) and - F(t) act during a head on collision. These forces change the linear momentum of both bodies.

Applying Newton's second law

Force $F = \dfrac{dP}{dt}$ and $dP = Fdt$ ……………………………………………. [17] where F (t) is a time varying force (see fig 3 (a))

Fig 3 a) magnitude of time varying force F(t) that action a body during a collision.

(b) Average Force Favg acting on a body over the time interval Δt(?)

$$\int_{Pi}^{Pf} d\vec{p} = \int_{ti}^{tf} \vec{F}(t)dt = \vec{P}_f - \vec{P}_i$$……………………………………………………[18]

Change in linear momentum = impulse J of collision = change in momentum

$$J = \int_{ti}^{tf} \vec{F}(t)dt$$……………………………………………………………………[19] impulse defined

$$\vec{P}_f - \vec{P}_i = \Delta \vec{P} = \vec{J}$$……………………………………………[20] impulse-linear momentum theorem

Momentum and Kinetic Energy in collisions

An elastic collision is the one which the kinetic energy of the system of colliding bodies is conserved. If some of the KE is transferred to other forms of energy eg thermal, sound or heat, KE of the system is not conserved and collision is inelastic.

Linear Momentum

In a closed, isolated system containing a collision, the linear momentum of each colliding body may change but the total linear momentum P of the system cannot change, whether the collision is elastic or inelastic-the law of conservation of linear momentum.

Inelastic collisions

Fig 4
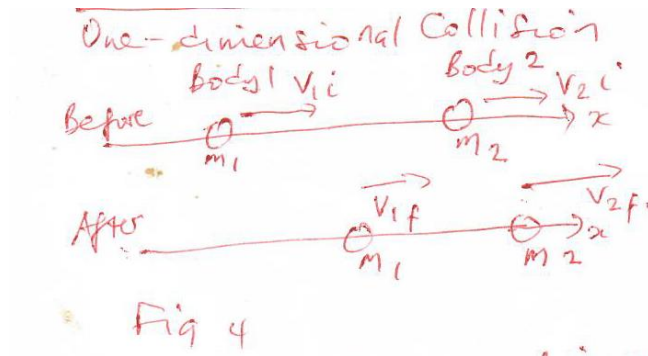
The law of conservation of linear momentum for this two body system is

(total momentum P_i before the collision) = (total momentum P_f after the collision)

$P_{1i} + P_{2i} = P_{1f} + P_{2f}$ conservation of linear momentum..............[21]

$m_1 v_{1i} + m_2 v_{2i} = m_{1f} v_{1f} + m_2 v_{2f}$...[22] P=mv

Completely Inelastic Collision
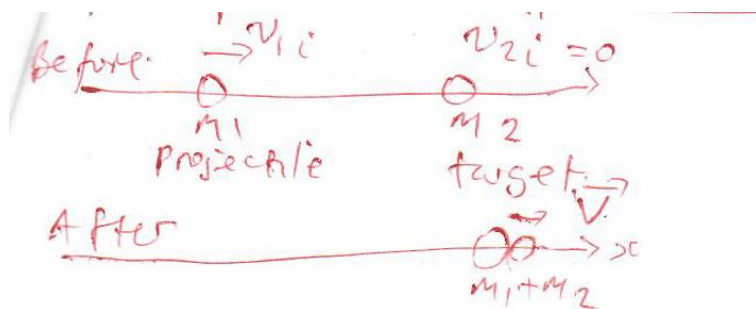
$m_1 v_{1i} + m_2 v_{2i} = (m_1 + m_2)V$..[23]

$or \ V = \frac{m_1}{m_1 + m_2} v_{1i}$...[24]

Velocity of Centre Mass

Fig 7.
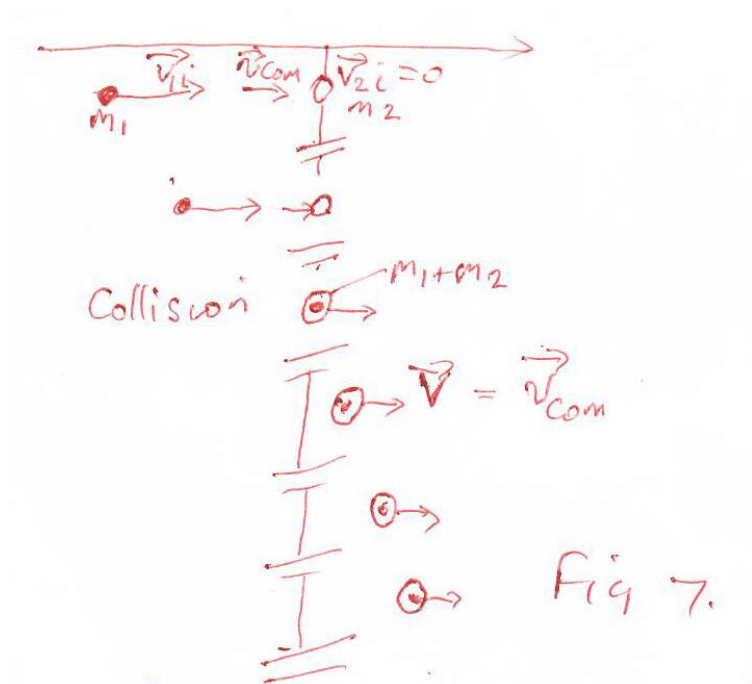
In a closed, isolated system velocity v_{com} of the centre mass of the system cannot be changed by a collision because with the system isolated, there is no net external force to change it. For a two body system.

$$P = Mv_{com} = (m_1 + m_2)v_{com}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[25]$$

$$= P_{1i} + P_{2i}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[26]$$

$$v_{com} = \frac{P}{(m_1+m_2)} = \frac{P_{1i}+P_{2i}}{m_1+m_2}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[27]$$

The RHS of this equation is a constant, and v_{com} has that same constant value before and after the collision.

Elastic Collisions in one Dimension

Stationary Target

In elastic collisions (total KE before the collision) = (total KE after the collision)

In an elastic collision, the KE of each colliding body may change ,but the total KE of the system does not change. Consider the elastic collision of masses m_1, and m_2, with m_2 initially stationary.
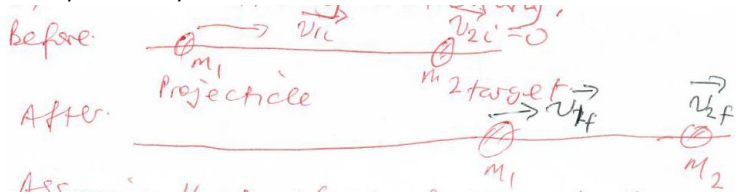
Assuming that above system is closed and isolated, the net linear momentum is conserved.

$m_1 v_{1i} = m_1 v_{1f} + m_2 v_{2f}$.........linear momentum...............[8]

Since collision is elastic, then the total KE is conserved

$\frac{1}{2} m_1 v_{1i}^2 = \frac{1}{2} m_1 v_{1f}^2 + \frac{1}{2} m_2 v_{2f}^2$ ... ...(K.E.)....................[9]

Subscripts i and f are initial and final respectively. From (8) rearranging

$m_1 (v_{1i} - v_{1f}) = m_2 v_{2f}$..[10]
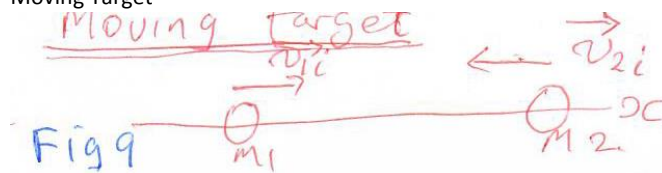
and from [9] $m_1 (v_{1i} - v_{1f})(v_{1i} + v_{1f}) = m_2 v_{2f}^2$................................[11]

From [10] and [11] $v_{1f} = \frac{m_1 - m_2}{m_1 + m_2} v_{1i}$..[12] and

$v_{2f} = \frac{2m_1}{m_1 + m_2} v_{1i}$..[13]

1. If masses are equal (12) and (13) reduce to $v_{1f} = 0$ and $v_{2f} = v_{1i}$
   It predicts that after a head on collision of bodies of equal masses, there will simply be an exchange of velocities.
2. For a massive target $m_2 \gg m_1$
   Equations (12) and(13) reduce to : $v_{1f} \approx -v_{1i}, and\ v_{2f} \approx (\frac{2m_1}{m_2}) v_{1i}$
3. For a massive projectile $m_1 \gg m_2$ the equations reduce to $v_{1f} \approx v_{1i}$
   and $v_{2f} \approx 2v_{1i}$..[14]

Moving Target

Fig 9

Conservation of linear momentum is

$$m_1 v_{1i} + m_2 v_{2i} = m_1 v_{1f} + m_2 v_{2f} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[15]$$

And conservation of of kinetic energy is

$$\frac{1}{2} m_1 v_{1i}^2 + \frac{1}{2} m_2 v_{2i}^2 = \frac{1}{2} m_1 v_{1f}^2 + \frac{1}{2} m_2 v_{2f}^2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[16]$$

To solve 15) and (16) we first rewrite

$$m_1 (v_{1i} - v_{1f}) = -m_2 (v_{2i} - v_{2f}) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[17]$$

And $m_1 (v_{1i} - v_{1f})(v_{1i} + v_{1f}) = -m_2 (v_{2i} - v_{2f})(v_{2i} + v_{2f}) \dots\dots\dots[18]$

$$v_{1f} = \frac{(m_1 - m_2)}{m_1 + m_2} v_{1i} + \frac{2m_2}{m_1 + m_2} v_{2i} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[19]$$

$$v_{2f} = \frac{2m_1}{m_1 + m_2} v_{1i} + \frac{m_2 - m_1}{m_1 + m_2} v_{2i} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[20]$$

Collisions in two dimensions

When collision is not head on, the bodies do not end up travelling along their initial axis. For such 2- dimensional collisions in a closed and isolated system the total linear momentum and total KE are conserved.

$$P_{1i} + P_{2i} = P_{1f} + P_{2f} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[21]$$

$$K_{1i} + K_{2i} = K_{1f} + K_{2f} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[22]$$

Fig

47
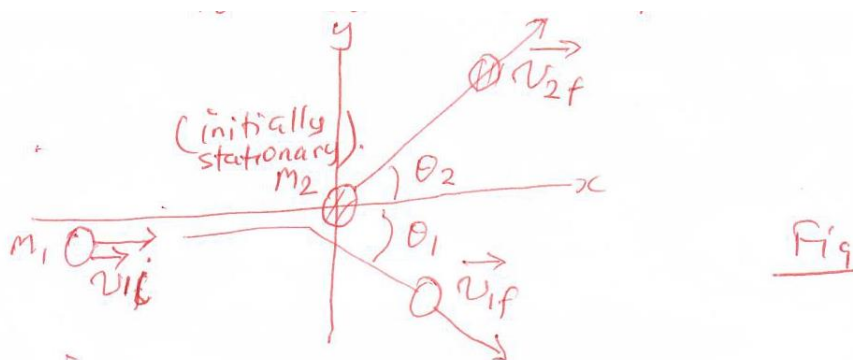
Rewriting (21) for components along x axis

$$m_1 v_{1i} = m_1 v_{1f} cos\theta_1 + m_2 v_{2f} cos\theta_2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[23]$$

And along y axis

$$0 = -m_1 v_{1i} sin\theta_1 + m_2 v_{2f} sin\theta_2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[24]$$

Equation (22) can be rewritten as $\frac{1}{2} m_1 v_{1i}^2 = \frac{1}{2} m_1 v_{1f}^2 + \frac{1}{2} m_2 v_{2f}^2$

WAVES

Types of Waves

1. Mechanical waves - they can exist only within a material medium eg water waves, sound waves ,seismic waves etc.

2.Electromagnetic waves -they don't require a material medium to exist eg visible, ultraviolet light, radio, television, x rays, microwaves. Light travels through a vacuum from stars to reach us.

3.Matter waves –these waves are associated with electrons ,protons and other fundamental particles even atoms and molecules.

4.Transverse and longitudinal waves –A transverse wave is one in which the particles of the transmitting medium vibrate in a direction perpendicular to the propagation of the wave. Eg a pulse moving along a string.

-A longitudinal wave is one in which the particles of a transmitting medium vibrate in a direction parallel to the propagation of the wave eg a sound wave.

Wavelength and frequency

For a transverse wave, the transverse displacement, y may be described as y = h (x,t).

Where x is the position of the element and t is the time.

As the wave progresses the displacement y at a position x and time t is

$y(x, t) = y_m sin(kx - wt)$ …………………………………………………………………………..(1)

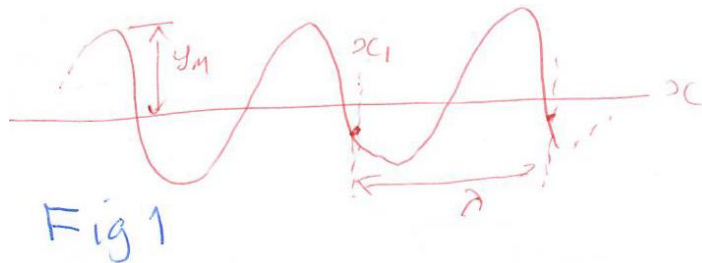y_m is the amplitude,

 k is angular wave number

sin (kx –wt) is the oscillating term

w is angular frequency

kx –wt is called the phase of the wave.

Fig 1

Amplitude and phase

The amplitude y_m of a wave is the magnitude of the maximum displacement of the elements from their equilibrium positions as the wave passes through them.

-The phase of the wave is the argument kx-wt of the sine in equation [1].

-As the wave sweeps through a string element at a particular position x, the phase changes linearly with time t.

Wavelength and angular wave number

The wavelength λ of a wave is the distance between repetitions of the shape of the wave.

From $y(x,t) = y_m\sin(kx - wt)$

At time $t = 0$, $y(x,0) = y_m \sin kx$ ………………………………………………………[2]

The displacement is the same at $x = x_1$ and $x = x_1 + \lambda$

$y_m\sin kx_1 = y_m\sin k(x_1 + \lambda)$

$= y_m\sin(kx_1 + k\lambda)$……………………………………………………………………………[3]

A sine function repeats itself when its angle (or argument) is increased by $2\pi \ rad$. ie

$k\lambda = 2\pi \ or \ k = \dfrac{2\pi}{\lambda}$…………………………………(angular wave number)……[4]

49

Period, angular frequency and frequency

Fig 2.
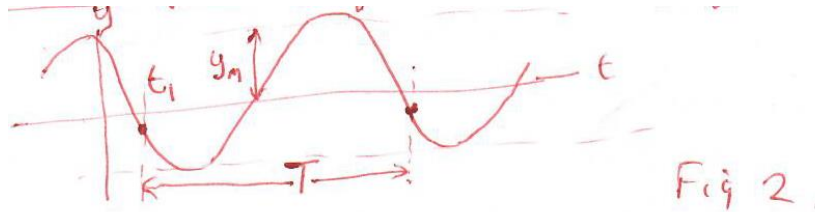
At x = 0, the sine element at that position moves up and down in a simple harmonic motion given by (1) with x = 0.

$y(0,t) = y_m\sin(-wt)$

$= -y_m\sin wt$................(x=0)..[5]

The period of oscillation T of a wave is the time any string element takes to move through one full oscillation.

Equating (4) to both ends of time interval yields

$-y_m\sin wt1 = -y_m\sin w(t_1+T)$

$= -y_m\sin(wt_1+wT)$..[6]

This can only be true if wT =2π or $\omega = \dfrac{2\pi}{T}$...[7](angular frequency).

The frequency f of the wave is $\dfrac{1}{T}$ $or\ f = \dfrac{1}{T} = \dfrac{2\pi}{\omega}$...............................[8]
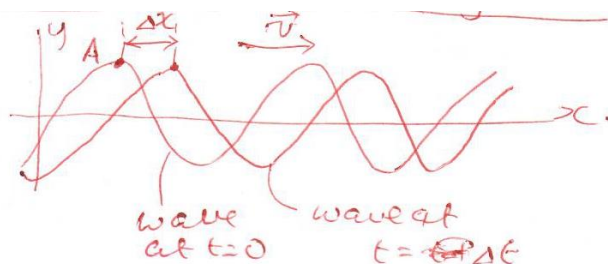
The speed of a travelling wave

Fig 3 Two snapshots of the wave at time and at t = Δt

Wave speed $v = \frac{\Delta x}{\Delta t}$ (or in differential limit $\frac{dx}{dt}$)

If A retains its displacement as it moves, the phase in [1] giving it that displacement must remain a constant: kx-wt = a constant…………………………………………………………..[9]

Taking a derivative of [9] yields: $k\frac{dx}{dt} - w = 0$ or $\frac{dx}{dt} = v = \frac{w}{k}$…………..[10]

$v = \frac{w}{k} = \frac{\lambda}{T} = \lambda f$…………………………..(wave speed)……………………………….[11]

Equation (1) describes a wave moving in the + ve direction of x. For a wave travelling in the opposite direction replace t by – t so that y(x,t) = y$_m$sin(kx+wt)……………..[12] and its

speed will be $\frac{dx}{dt} = -\frac{w}{k}$……………………………………………………………….[13]

Wave speed on a stretched string

The mass and elasticity properties of a medium determine the speed of the wave in that

medium. $v = \sqrt{\frac{\tau}{\mu}}$ ……………………………………………………………….[14] where τ is

tension and μ is linear density(mass per unit length)

The speed of a wave along a stretched ideal string depends only on the tension and linear density of the string and not on the frequency of the wave.
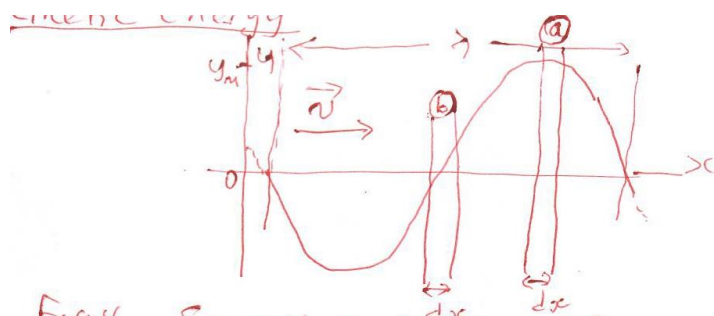
Energy and power of a travelling wave

Kinetic energy

Fig 4. Snap shot of a travelling wave at time t = 0. String element a is at displacement y_m. String element b is at displacement y = 0.

Consider an element of string of mass dm, oscillating transversely in simple harmonic motion. As the wave passes through it, it has kinetic energy associated with its transverse velocity u. The transverse velocity and hence KE is a maximum at y =0 and zero at y = y_m. The kinetic energy dK associated with a string element of mass dm is given by

$$dK = \frac{1}{2}dmu^2 \dotfill [15]$$

$$u = \frac{dy}{dt} = -wy_m\cos(kx - wt) \dotfill [16]$$

Using dm = μdx, $dK = \frac{1}{2}(\mu dx)(-wy_m)^2\cos^2(kx - wt) \dotfill [17]$

$$\frac{dK}{dt} = \frac{1}{2}\mu vw^2 y_m^2\cos^2(kx - wt) \dotfill [18]$$
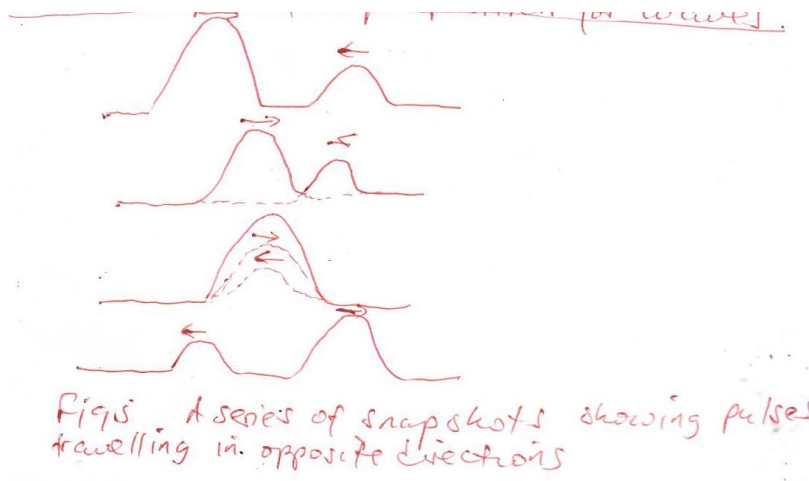
The average at which kinetic energy is transported is:

$$\left(\frac{dK}{dt}\right)_{avg} = \frac{1}{2}\mu vw^2 y_m^2[cos^2(kx - wt)]_{avg}.$$

$$= \frac{1}{4}\mu vw^2 y_m^2 \dotfill [19]$$

The average power $P_{avg} = 2\left(\frac{dK}{dt}\right)_{avg} \dotfill [20]$

$$= \frac{1}{2}\mu vw^2 y_m^2 \dotfill [21]$$

The principle of superposition for waves

Fig15. A series of snapshots showing pulses travelling in opposite directions

The displacement of the string when the waves overlap is the algebraic sum.

$$y'(x,t) = y_1(x,t) + y_2(x,t) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [22]$$

Overlapping waves algebraically add to produce a resultant wave (or net wave).

This is an example of the principle of superposition which says that when several effects occur simultaneously, their net effect is the sum of the individual effects.

Overlapping waves do not in any way alter the travel of each other.

Interference of Waves

When two waves meet the resultant wave depends on the extent to which the waves are in phase. Let the two waves be given by :

$$y_1(x,t) = y_m \sin(kx - wt) \text{ and}$$

$$y_2(x,t) = y_m \sin(kx - wt + \Phi).$$

The two waves have the same angular frequency w, same angular wave number k, travel with the same speed in the same direction but only differ in phase.

Displacement, $y'(x,t) = y_1(x,t) + y_2(x,t)$

$$= y_m \sin(kx - wt) + y_m \sin(kx - wt + \Phi).$$

$$= \left[2y_m \cos \frac{1}{2}\Phi \right] \sin(kx - wt + \Phi) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [23]$$

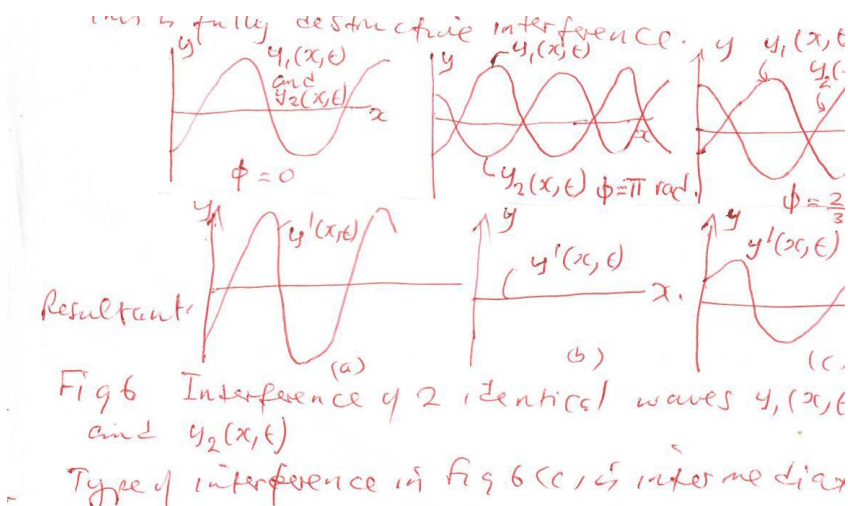$\left[2y_m cos\frac{1}{2}\Phi\right]$ is the amplitude and $sin(kx - wt + \Phi)$ is the oscillating term.

If two sinusoidal waves of the same amplitude and wavelength travel in the same direction along a stretched string, they interfere to produce a resultant sinusoidal wave travelling in that direction. If the two waves are in phase and $\Phi = 0$

$$y'(x,t) = 2y_m sin(kx - wt)\text{.......................................}(\Phi = 0)\text{.....................}[24]$$

Interference that produces the greatest possible amplitude is called fully constructive interference.

If $\Phi = \pi\,rad$ (or 180^0) the interfering waves are exactly out of phase, the amplitude of resultant wave is zero. $y'(x,t) = 0 \text{ ... ...}(\Phi = \pi\,rad)\text{............}[25]$

This is fully destructive interference.

Fig 6   Interference of 2 identical waves $y_1(x,t)$ and $y_2(x,t)$

Type of interference in fig 6(c) is intermediar

Phasors

A phasor is a vector that has a magnitude equal to the amplitude of the wave and that rotates round an origin.

The angular speed of the phasor is equal to the angular frequency w of the wave.eg wave
$y_1(x,t)=y_{m1}\,sin(kx- wt)$ ....[26] is represented by phasor in Fig 7 (a)

The magnitude of the phasor is the amplitude y_m of the wave. As phasor rotates about orig

(9)

a)   b)   c)

Fig 7
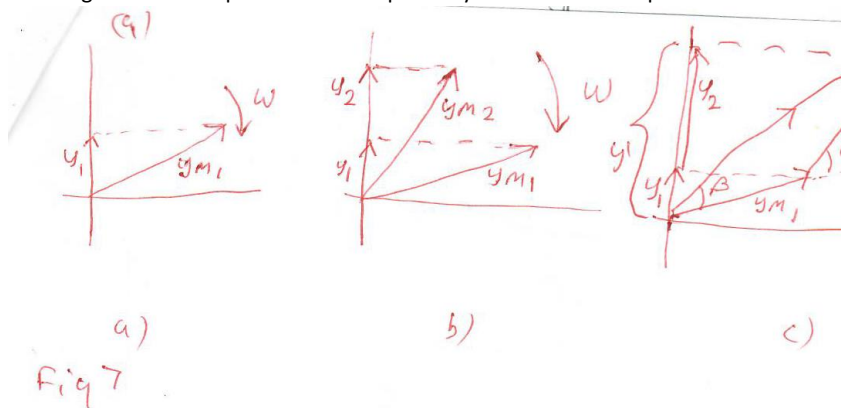
in its

projection y on the vertical axis varies sinusoidally from a maximum of y_{m1} through zero to a minimum of $-y_{m1}$ and then back to y_{m1}.

When two waves travel along the same string in the same direction , they and their resultant can be represented on a phasor diagram. The phasor in Fig 7(b)

$y_2(x,t) = y_{m2}\sin(kx-wt+\Phi)$ is shifted from the first by phase constant Φ. If Φ is +ve then the phasor for wave 2 lags the phasor for wave 1 as they rotate(Fig 7 (c)). . If Φ is -ve then the phasor for wave 2 leads the phasor for wave 1.

Because y_1 and y_2 have same k and w then their resultant is in the form :

$$y'(x,t) = y'_m\sin(kx - wt + \beta)\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[27]$$

Where y'_m is the amplitude of the resultant and β is its phase constant.

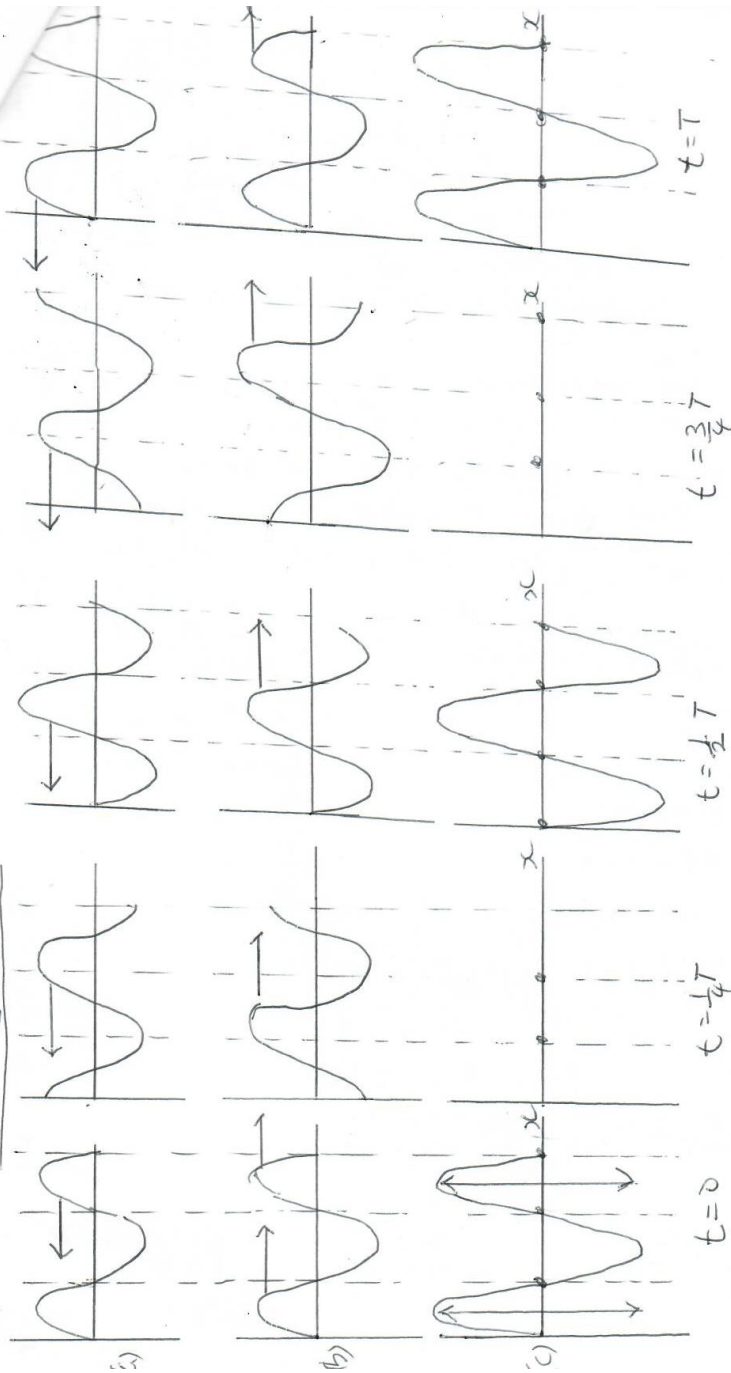We can use phasors to combine waves even if their amplitudes are different.

Standing Waves

(a) travelling to the left
(b) "   " "   right
(c) the resultant wave.

Fig 8. Snapshots of waves.

t=0   t=¼T   t=½T   t=¾T   t=T

(a)
(b)
(c)

Standing waves

Fig 8a, b, c show two waves one travelling to the right , the other travelling to the left on the same stretched string and their resultant after applying the principle of superposition graphically.

There are places called nodes along the string where the string never moves (marked by dots in resultant wave.)

Halfway between successive nodes are antinodes, where the amplitude of the resultant wave is maximum.

Wave patterns in (c) are called standing waves because the wave patterns do not move left or right, the locations of maxima and minima do not change.

If 2 sinusoidal waves of the same amplitude and wavelength travel in opposite directions along a stretched string, their interference with each other produces a standing wave.

If the two combining waves are : $y_1(x,t) = y_m \sin(kx - wt) \ldots \ldots \ldots \ldots \ldots \ldots [28]$ and $y_2(x,t) = y_m \sin(kx + wt) \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [29]$ then the principle of superposition for the combined wave is:

Displacement, $y'(x,t) = y_1(x,t) + y_2(x,t)$

$= y_m \sin(kx - wt) + y_m \sin(kx + wt).$

$= [2y_m \sin kx] \cos wt \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [30]$ where

$[2y_m \sin kx]$ is the amplitude at position x and $\cos wt$ is the oscillating term.

This is an equation of a standing wave. The amplitude varies with position. It is zero for values of kx = 0 = nπ where n = 0 ,1, 2- - -$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$[31]

Substituting for k = 2π/λ in (31) and rearranging yields $x = \dfrac{n\lambda}{2}$, n=0, 1, 2…….[32](nodes)

Amplitude of standing wave of (30) has a maximum value of $2y_m$ occurring for values of kx that give sinkx = 1, ie. $kx = \dfrac{1}{2}\pi, \dfrac{3}{2}\pi, \dfrac{5}{2}\pi \ldots \ldots \ldots = (n + \dfrac{1}{2})\pi$ for n= 0, 1, 2………..[33]

Substituting k = $2\pi/\lambda$ in [33] and rearranging yields:

$x = (n + \dfrac{1}{2})\dfrac{\lambda}{2}$, for n = 0, 1, 2………………(antinodes)…………………………………………[34]
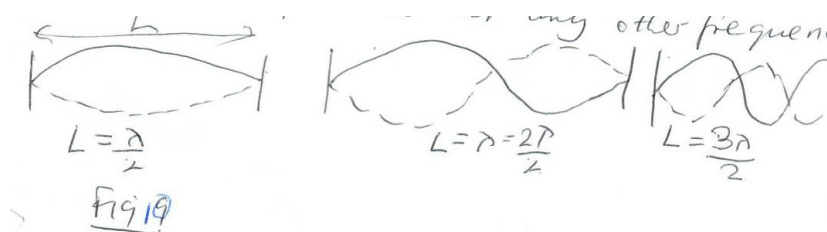
Standing waves and resonance

When a continuous sinusoidal wave travelling to the right is reflected, the on-going and reflected wave overlap and form a standing wave pattern .(Fig 9)
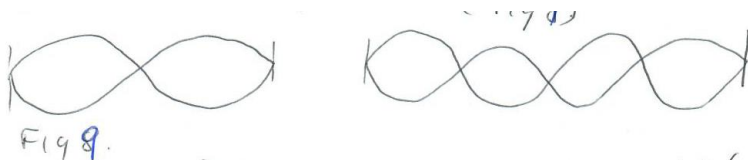
Fig 9.

For certain frequencies, the interference produces a standing wave pattern (oscillation mode) Fig 9.

Such a standing wave is said to be produced at resonance, and the string is said to resonate at these resonance frequencies. A standing wave pattern is not produced at any other frequencies.

Fig 19

A standing wave can be set up on a string of length L by a wavelength equal to :

$$\lambda = 2L, 2\frac{L}{2}, 2\frac{L}{3} etc$$

$\lambda = \frac{2L}{n} \ for \ n = 1, 2, 3 \dots \dots$ ....[35]

The resonant frequencies corresponding to these wavelengths are:

$f = \frac{v}{\lambda} = n\frac{v}{2L} \ for \ n = 1, 2, 3 \dots \dots \dots \dots$ ....[36]

SOUND WAVES

Mechanical waves are waves that require a material medium to exist eg transverse and longitudinal where oscillations are perpendicular and parallel to direction of wave travel respectively. Sound is a longitudinal wave.

Fig 1

A sound wave travelling through 3-dimensional medium from source S.
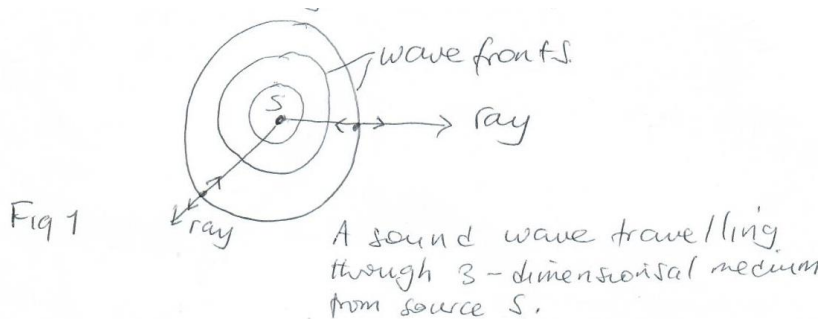
S above represents a point source, that emits sound in all directions. Wave fronts and rays indicate the direction of travel and the spread of sound waves.

Wave fronts are surfaces over which the oscillations of the air due to the sound wave have the same value.

Rays are directed lines perpendicular to the wave fronts that indicate the direction of travel the wavefronts.

Near the source wave fronts are spherical but far from the sources as curvature decreases the wave fronts become planar.

The speed of sound

The speed of any mechanical wave depends on both an inertial property of the medium (to store kinetic energy) and the elastic property of the medium (to store potential energy)

The speed of a transverse wave along a stretched string is :

$$v = \sqrt{\frac{\tau}{\mu}} = \sqrt{\frac{elastic\ property}{inertial\ property}} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[1]$$

Where τ is tension of the string and μ is the linear density of the string.

If the medium is air and wave is longitudinal the inertial property corresponding to μ is the volume density ρ of the air.

The elastic property corresponding to τ is the bulk modulus B, which is defined as

$$B = -\frac{\Delta P}{\Delta V/V} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots[2]$$

the change in pressure to fractional change in volume .

$$v = \sqrt{\frac{B}{\rho}} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(\text{speed of sound})\dots\dots\dots\dots[3]$$
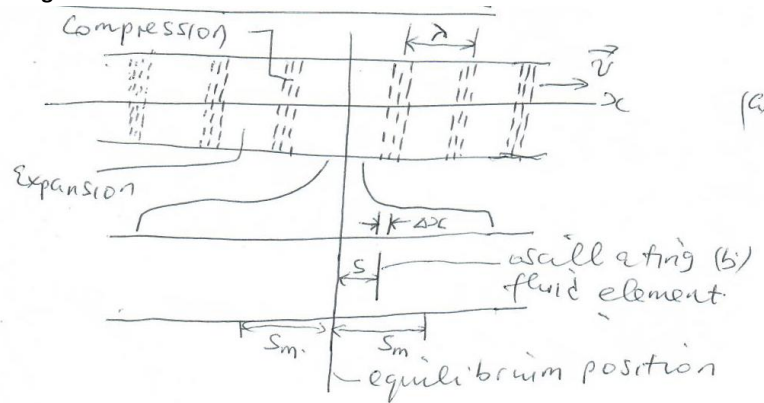
Travelling sound waves

Fig 2ᵃ

Fig 2(a) shows a sound wave, travelling through a long air filled tube with speed v, consisting of a moving, periodic pattern of expansions and compressions of the air .

(b) A horizontally expanded view of a piece of the tube. The wave is produced by sinusoidally moving a piston to and fro.

Consider a thin element of air of thickness Δx at position x along the tube. As the wave travels through x, the element of air oscillates left and right in simple harmonic motion about its equilibrium position.

Because air components oscillate parallel to the x axis,

displacement $s(x,t) = s_m\cos(kx - wt)$..[4] s_m is amplitude.

And pressure variation $\Delta p(x,t) = \Delta p_m\sin(kx - wt)$....................[5] Δp_m is pressure variation amplitude.

As the wave moves, the air pressure at any position x varies sinusoidally as in (5).

A −ve value in (5) corresponds to an expansion and a +ve value corresponds to a compression.

$\Delta p_m = (v\rho w)s_m$..[6]

Interference

Interference

S_1 L_1 → P

S_2 → L_2

Fig 3

Two sound sources S_1 and S_2 emit spherical sound waves in phase and identical wav
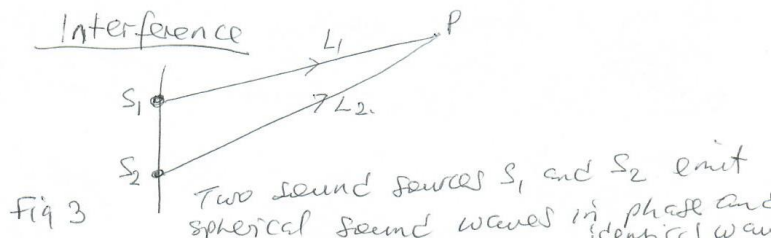
Interference.

New position of wavefront at time $t = \Delta t$

wavefront at $t = 0$

Fig 1; The propagation of a plane wave in a vacuum, as portrayed by Huygens' principle

All points on a wavefront serve as a point source of spherical secondary wavelets. After a time t, the new position of the wavelength will be ~~bent~~ that of a surface tangent to these secondary wavelets.

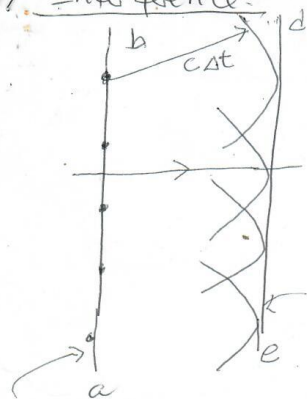Dots on ab serve as point sources of spherical waves. The radii of the waves grow by $c\Delta t$ after time Δt where c is speed of light. The plane tangent to these wavelets represents the new wavefront de. de is parallel to ab a perpendicular distance $c\Delta t$ from it.

The law of refraction

n_1

n_2
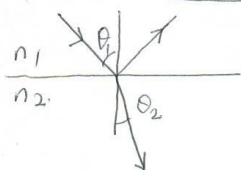
$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad ---(1)$$

Snell's law

If the 2 waves travel equal distances ie ($L_1 = L_2$) to reach P then they will be in phase.

If however the paths travelled are different they may not be in phase at P.

Their phase difference depends on their path length difference $\Delta L = |L_2 - L_1|$

Phase difference Φ is related to the path length difference by : $\dfrac{\Phi}{2\pi} = \dfrac{\Delta L}{\lambda}$..........[7] as a path difference of 2π corresponds to one wavelength. Therefore $\Phi = \dfrac{\Delta L}{\lambda} 2\pi$............[8]

$\Phi = m(2\pi)$ for m= 0, 1, 2.......fully constructive interference...............................[9]

$\Phi = (2m + 1)\pi$ for m=0, 1, 2.....fully destructive interference.............................[10]

From [8] and [9] $\dfrac{\Delta L}{\lambda}$= 0, 1, 2......fully constructive interference...........................[11]

From [10] $\dfrac{\Delta L}{\lambda} = 0.5, 1.5, 2.5$...[12]

Intensity and sound level

The intensity I of a sound wave at a surface is the average rate per unit area at which energy is transferred by the wave through or onto the surface.

$$I = \frac{P}{A}$$...[13]

P is time rate of energy transfer (the power) of the sound wave and A is the area of surface intercepting the sound.

$$I = \frac{1}{2}\rho v w^2 s_m^2$$ ....[14]

Where ρ is density of air, v is speed of wave.

Fig4 A sound source S emits sound wave uniformly in all directions.
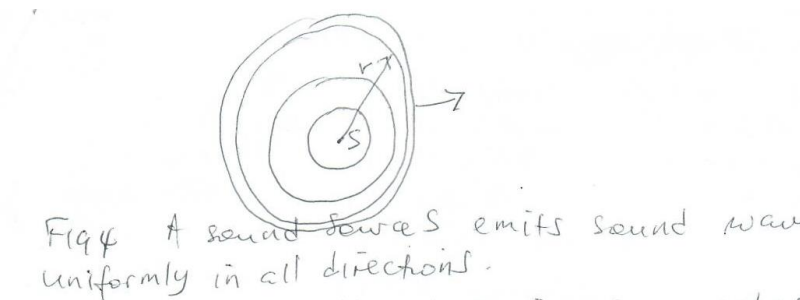
The waves pass through an imaginary sphere of radius r that is centred on S.

$$I = \frac{P}{A} = \frac{P}{4\pi r^2}$$

The intensity of a sound wave decreases with the squared distance r from the source.

The Doppler effect

If a police car parked by the road side sounds a 1000 Hz siren, if you are also parked by the road side you will hear the same frequency. If you are driving towards the police car at 120 km / hr you will hear a higher frequency of 1096 Hz but if you are driving away you will hear a lower frequency of 904 Hz a decrease of 96Hz.

These motion related changes in frequencies are examples of the Doppler effect. This effect also holds for electromagnetic waves.

If the source and detector move either directly toward or away from each other, at speeds less than the speeds of sound

$$f' = f \frac{v \pm v_D}{v \pm v_S}$$ -Doppler effect..[16]

Where v is the speed of sound through air

v_D is the detector's speed relative to the air

v_S is the source's speed relative to the air

f' is the detected frequency, f is emitted frequency

When the motion of source and detector is towards each other, the sign on its speed must give an upward shift in frequency.

When the motion of detector or source is away from the other the sign on its speed must give a downward shift in frequency.

INTERFERENCE
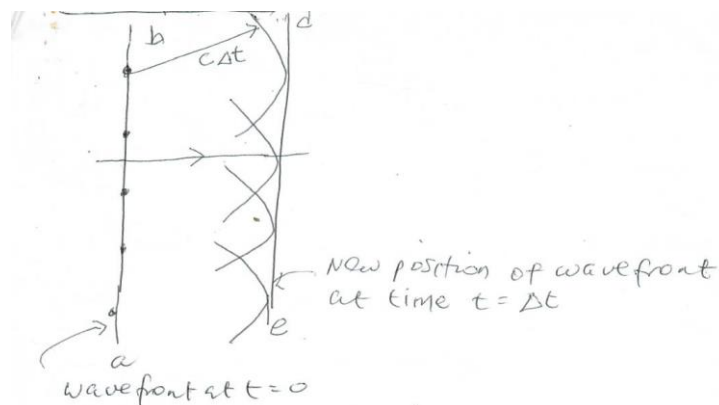
Fig 1; The propagation of a plane wave in a vacuum, as portrayed by Huygens' principle

All points on a wave front serve as point sources of spherical secondary wavelets.

Dots on ab serve as point sources of spherical waves. The radii of the waves grow by ct after time t where c is speed of light. The plane tangent to these wavelets represents the new wave front de. De is parallel to ab a perpendicular distance ct from it.
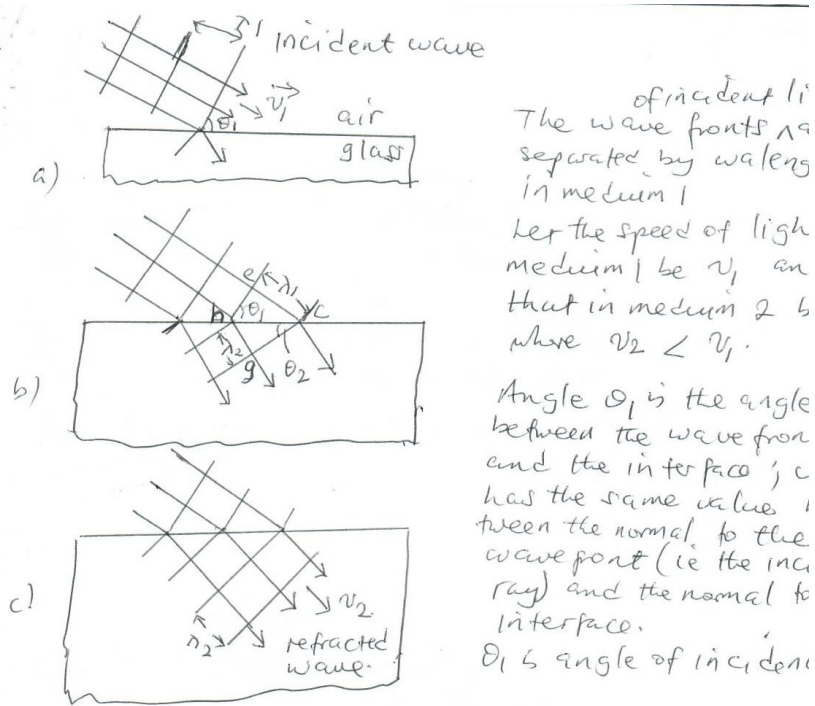
The law of refraction

In the figure, handwritten annotations:

a) incident wave, $\vec{v_1}$, air, glass, θ_1

b) $e \leftarrow \lambda_1$, θ_1, c, h, λ_2, g, θ_2

c) v_2, n_2, refracted wave.

Handwritten notes (right side):

of incident li'
The wave fronts ag
separated by waleng
in medium 1
Let the speed of ligh
medium 1 be v_1 an
that in medium 2 b
where $v_2 < v_1$.

Angle θ_1 is the angle
between the wave fron
and the interface; c
has the same values
tween the normal to the
wave front (ie the inc
ray) and the normal to
interface.
θ_1 is angle of incident

Fig 2, Three stages in the refraction of several e fronts at a plane interface air (medium 1) and glass (medium 2).

In (b) As the wave moves into the glass, a Huygens wavelet at point e will expand to pass through point c, at a distance of λ_1 from point e.

The time interval for this expansion is $\dfrac{\lambda_1}{v_1}$

At the same time a Huygens wavelet at point h will expand to pass through point g, at a reduced speed v_2 with wavelength λ_2. Time interval must be $\dfrac{\lambda_2}{v_2}$.

Therefore $\dfrac{\lambda_1}{v_1} = \dfrac{\lambda_2}{v_2}$ or $\dfrac{\lambda_1}{\lambda_2} = \dfrac{v_1}{v_2}$...[2]

Equation 2 shows that the wavelengths of light in two media are proportional to the speeds of light in the media.

By Huygens' principle the refracted wave fronts must be tangent to an arc of radius λ_2 centred on h, at g.

It must also be tangent to arc of radius λ_1 centred on e, at c.

66

The angle between the refracted wave front and interface θ_2 is the angle of refraction.

For right angles hce and hcg we may write $\sin\theta_1 = \dfrac{\lambda_1}{hc}$ (for Δhce)...................[3]

$\sin\theta_2 = \dfrac{\lambda_2}{hc}$ (for Δhcg)..[4]

$[3] \div [4]\ yields\ \dfrac{\sin\theta_1}{\sin\theta_2} = \dfrac{\lambda_1}{\lambda_2} = \dfrac{v_1}{v_2}$..[5]

If we define the index of refraction n as ratio of speed of light in vacuo to speed of light v in medium $n = \dfrac{c}{v}$...[6]

Then $n_1 = \dfrac{c}{v_1}$ and $n_2 = \dfrac{c}{v_2}$..[7]

From (5) and (7) we find $\dfrac{\sin\theta_1}{\sin\theta_2} = \dfrac{c/n_1}{c/n_2} = \dfrac{n_2}{n_1}$...[8]

Or $n_1 \sin\theta_1 = n_2 \sin\theta_2$........Snell's law of refraction...[9]

Wavelength and jndex of refraction

The wavelength of light changes as the speed of light changes when light crosses an interface from one medium into another.

The speed of light in any medium depends on the index of refraction of the medium (6).

Thus the wavelength also depends on the angle of refraction.

(5) can be rewritten as $\lambda_n = \lambda \dfrac{v}{c}$...[10]

Where λ_n is wavelength in medium

λ is wavelength in vacuum

v is the speed in medium
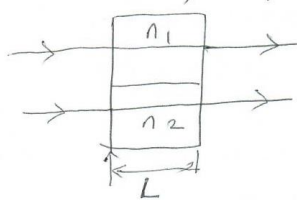
c is the speed in vacuum

Using (6) to substitute $\dfrac{1}{n}$ for $\dfrac{v}{c}$ $then\ yields\ \lambda_n = \dfrac{\lambda}{n}$...[11]

And since v = fλ we can write $f_n = \dfrac{v}{\lambda_n} = \dfrac{c/n}{\lambda/n} = \dfrac{c}{\lambda} = f$

Although speed and wavelength of light in medium changes, frequency remains the same.

Fig 3

Two light rays travelling through two media having different indices of refraction

The two waves of identical wavelengths pass through different media with refractive indices n_1 and n_2 of length L.

Initially the incident waves are in phase in air ($n \approx 1$).

After emerging from the two media they will have the same wavelength (their wavelength in air). However because their wavelengths differed in the two media, the two waves may no longer be in phase.

The phase difference between two light waves can change if the waves travel through different materials having different indices of refraction. The phase difference can determine how the light waves will interfere if they reach some common point.

To determine the new phase difference we first find the number of wavelengths in each medium. No. of wavelengths in medium 1: $N_1 = \dfrac{L}{\lambda_{n1}} = \dfrac{Ln_1}{\lambda}$...............................[12] since

$\lambda_{n1} = \lambda/n_1$, No. of wavelengths in medium 2: $N_2 = \dfrac{L}{\lambda_{n2}} = \dfrac{Ln_2}{\lambda}$....................[13]

Assuming $n_2 > n_1$ $N_2 - N_1 = \dfrac{Ln_2}{\lambda} - \dfrac{Ln_1}{\lambda} = \dfrac{L}{\lambda}(n_2 - n_1)$..........................[14]

If (14) tells us that $N_2 - N_1 = 45{,}6$ wavelengths. However a shift of an integer number of wavelengths such as 45 would put the waves back in phase so it is only the decimal fraction 0,6 which is important. A phase difference of 45,6 wavelengths is equivalent to an effective phase difference of 0,6 wavelengths.

Young's interference experiment

Fig 1
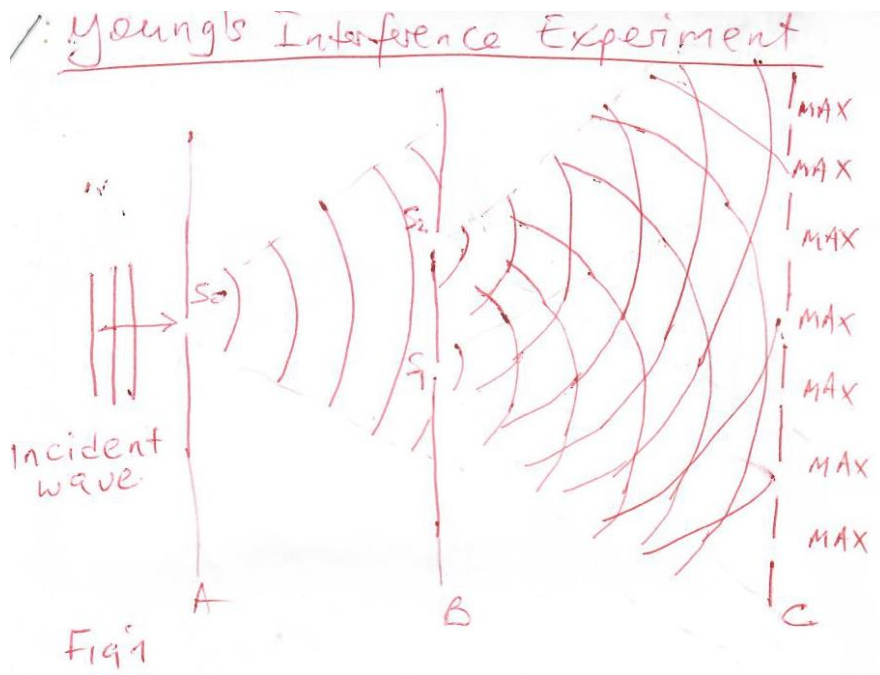
Light from a distant monochromatic source illuminates slit S_O in screen A. The emerging light then spreads via diffraction to illuminate two slits S_1 and S_2 in screen B. Diffraction of the light by the 2 slits sends overlapping circular waves into the region beyond screen B, where the waves from one slit interfere with the waves from the other slit. Points of constructive interference and destructive interference form bright bands (fringes) and dark bands (fringes) respectively on screen C. The pattern of bright and dark fringes is called an interference pattern.

Locating the fringes

A plane wave of monochromatic light is incident on 2 slits S_1 and S_2 on screen B; the light diffracts through the slits and produces an interference pattern on screen C.

Fig 2 — path length difference ∆L
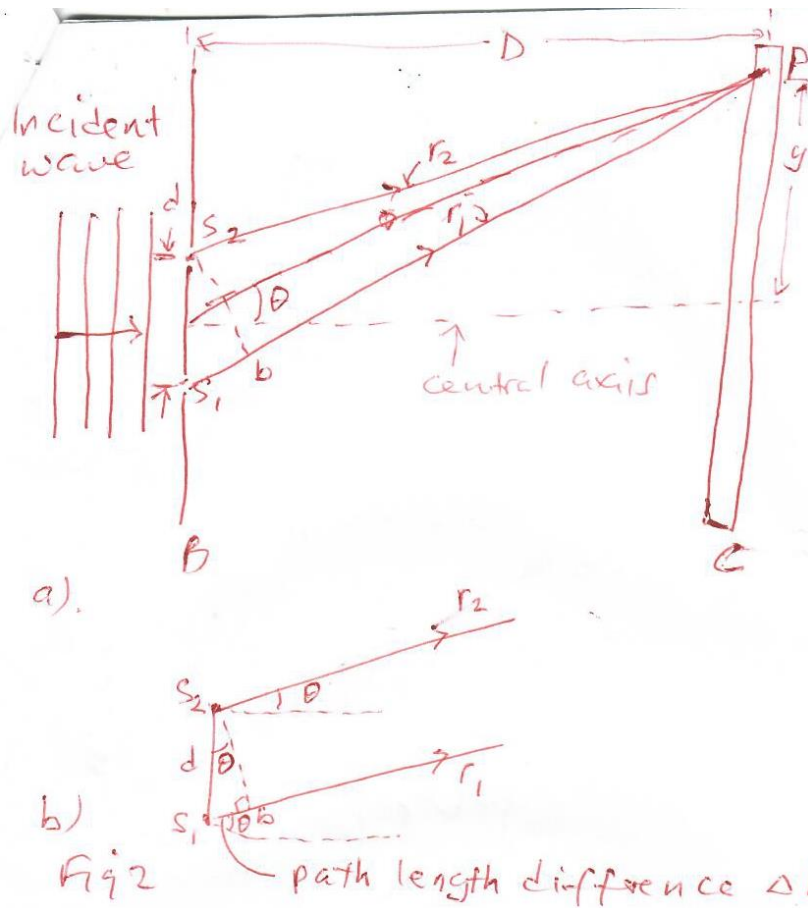
The waves of ray r_1 and r_2 are in phase when they pass through the slits S_1 and S_2. The two waves travel different distances to reach P. The phase difference between 2 waves can change if the waves travel paths of different lengths.

If ∆L is an integer number of wavelengths the waves arrive in phase at P, and interfere fully constructively making P part of a bright fringe.

If ∆L is an odd multiple of half wavelengths the waves arrive out of phase at P, and they interfere fully destructively making P part of a dark fringe.

What appears at each point on the viewing screen in a Young's double -slit interference experiment is determined by the path length difference ∆L of the rays reaching that point.

If D≫d in (b) then we can approximate rays r_1 and r_2 as being parallel to each other and at angle θ to the central axis.

We can also approximate S_1S_2b to be right angled and $S_1\hat{S}_2b$ to be θ.

Therefore $sin\theta = \frac{\Delta L}{d}$ and $\Delta L = dsin\theta$..[1]

For a bright fringe, ΔL = 0 or an integer multiple of wavelengths

ΔL = dsinθ = mλ, m = 0, 1, 2..[2]

For a dark fringe, ΔL is an odd multiple of half a wavelength

$\Delta L = dsin\theta = \left(m + \frac{1}{2}\right)\lambda$ $m = 0, 1, 2$..[3]

The central maximum is the point at which waves arriving from the 2 slits have a path difference ΔL = 0, hence zero phase difference.

For m=2 bright fringes are at an angle $\theta = sin^{-1}\left(\frac{2\lambda}{d}\right)$ above and below the central axis. These are called 2nd order fringes (m=2).

From [3] for m=1 dark fringes are at an angle θ = $sin^{-1}\left(\frac{1.5\lambda}{d}\right)$ above and below the central axis. These are second order dark fringes. The first order dark fringes occur when m = 0.

For the mth bright fringe dsinθ = mλ

Assuming θ to be small then sinθ = tanθ = θ in radians

But $= \frac{y_m}{D}$,$\therefore d\frac{y_m}{D} = m\lambda$

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

For the next further out maximum $y_{m+1} = \frac{(m+1)\lambda D}{d}$

$$\therefore Fringe\ separation\ \Delta y = y_{m+1} - y_m$$

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

$$= \frac{\lambda D}{d} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. [4]$$

The Michelson Interferometer

An interferometer is a device that can be used to measure lengths or changes in length with great accuracy by means of interference fringes.
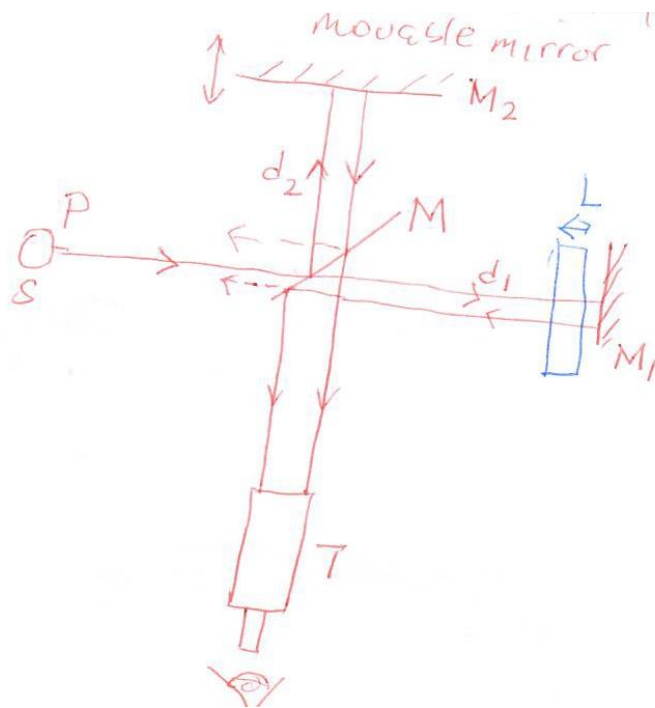

Fig 3

Consider light from point P on extended source S in Fig 3 and encounters beam splitter M. A beam splitter is a mirror that transmits half the incident light and reflects the other half.

At M the light divides into 2 waves, one proceeds by transmission towards mirror M_1, the other proceeds by reflection towards mirror M_2.

The waves are entirely reflected at the mirrors and are sent back along their directions of incidence, each wave eventually entering telescope T. The observer sees a pattern of curved or a pattern of approximately straight interference fringes.

The path length difference for the 2 waves when they recombine at telescope is $2d_2 - 2d_1$.

If mirror M_2 is moved by a distance $\frac{1}{2}\lambda$, the path difference is changed by λ and the fringe pattern will shift by one fringe.

Similarly moving mirror M_2 by $\frac{1}{4}\lambda$ causes a shift by half a fringe (each dark stripe shifts to where the adjacent white stripe was).

A shift in the fringe pattern can also be caused by the insertion of a thin transparent material into the optical paths of one of the mirrors say M_1. If the material has thickness L

and index of refraction n, then the number of wavelengths along the light's to and fro path through the material is from . $N_1 = \frac{L}{\lambda_{n1}} = L\frac{n_1}{\lambda}, N_m = \frac{2L}{\lambda_n} = \frac{2Ln}{\lambda}$..........................[5]

The number of wavelengths in the same thickness 2L of air before insertion of material is

$N_a = \frac{2L}{\lambda}$...[6]

When material is inserted the light returned by mirror M1 undergoes phase change of

$N_m - N_a = \frac{2Ln}{\lambda} - \frac{2L}{\lambda} = \frac{2L(n-1)}{\lambda}$...[7]

Thus by counting the number of fringes through which the material causes the pattern to shift, and substituting that for N_m-N_a you can determine the thickness L of the material in terms of λ.