

INFORMATION & COMMUNICATION TECHNOLOGY SYSTEM MAINTENANCE

1st Semester

TRADE THEORY

SECTOR: IT & ITES



GOVERNMENT OF INDIA
MINISTRY OF LABOUR & EMPLOYMENT
DIRECTORATE GENERAL OF EMPLOYMENT & TRAINING



**NATIONAL INSTRUCTIONAL
MEDIA INSTITUTE, CHENNAI**

Post Box No. 3142, CTI Campus, Guindy, Chennai - 600 032

Sector : IT & ITES

Duration : 2 - Years

**Trade : Information & Communication Technology System Maintenance -
Trade Theory**

Copyright © 2014 National Instructional Media Institute, Chennai

First Edition : November 2014,

Copies : 1,000

Rs.320/-



All rights reserved.

No part of this publication can be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system, without permission in writing from the National Instructional Media Institute, Chennai.

Published by:

**NATIONAL INSTRUCTIONAL MEDIA INSTITUTE
P. B. No.3142, CTI Campus, Guindy Industrial Estate,
Guindy, Chennai - 600 032.**

Phone: 044 - 2250 0248, 2250 0657

Fax : 91 - 44 - 2250 0791

email : nimichennai@vsnl.net , nimi_bsnl@dataone.in

Website: www.nimi.gov.in

FOREWORD

The National Instructional Media Institute (NIMI) is an autonomous body under the Directorate General of Employment and Training (DGE&T) Ministry of Labour and Employment has been developing, producing and disseminating Instructional Media Packages (IMPs) are extensively used in the Industrial Training Institutes/Training centres in Industries to impart practical training and develop work-skills for the trainees and the trainers

The Ministry of Labour & Employment constituted Mentor Councils (MCs) to revamp courses run / to be run under National Council of Vocational Training (NCVT) in 25 sectors. The MCs have representatives from thought leaders among various stakeholders viz. one of the top ten industries in the sector innovative entrepreneurs who have proved to be game-changers, academic/professional institutions (IITs etc.), experts from field institutes of DGE &T, champion ITIs for each of the sectors and experts in delivering education and training through modern methods like through use of IT, distance education etc. The technical support to the MCs is provided by Central Staff Training and Research Institute (CSTARI), Kolkata and National Instructional Media Institute (NIMI), Chennai. Some of the MCs are also supported by sector-wise Core Groups which were created internally in the Ministry (in 11 sectors).

A Steering Committee to provide overall coordination and guidance to Mentor Councils has also been constituted and has representation from the MCs, Chair positions to be endowed by the Ministry, trade unions, and experts on distance education and training. The MCs are mandated to work towards revamping/suggesting new courses, improving assessment systems, overall learning etc. for subjects under the purview of the NCVT.

Accordingly NIMI with the support and assistance of MC has developed **Information & Communication Technology System Maintenance Trade Theory 1st Semester in IT & ITES Sector** to enhance the employability of ITI trainees across the country and also to meet the industry requirement.

I have no doubt that the trainees and trainers of ITIs & Training centres in industries will derive maximum benefit from these books and that NIMI's effort will go a long way in improvement of Vocational Training.

I complement Director, Mentor Council members, Media Development Committee (MDC) members and staff of NIMI for their dedicated and invaluable contribution in bringing out this publication.



ALOK KUMAR, I.A.S.,
Director General of Employment &
Training/ Joint Secretary
Ministry of Labour and Employment
Government of India

New Delhi - 110 001

PREFACE

This National Instructional Media Institute (NIMI) was set up at Chennai by the Directorate General of Employment and Training (DGE&T) Ministry of Labour and Employment, Government of India with technical assistance from the Govt. of the Federal Republic of Germany. The prime objective of this institute is to develop and disseminate instructional materials for various trades as per the prescribed syllabi under the Craftsmen and Apprenticeship Training Schemes.

The instructional materials are developed and produced in the form of Instructional Media Packages (IMPs). An IMP consists of Trade Theory book, Trade Practical book, Test and Assignment book, Instructor guide, Wall Charts and Transparencies.

Hon'ble Union Minister of Finance during the budget speech 2014-2015 mentioned about developing **Skill India** and made the following announcement

"A national multi-skill programme called Skill India is proposed to be launched. It would skill the youth with an emphasis on employability and entrepreneur skills. It will also provide training and support for traditional professions like welders, carpenters, cobblers, masons, blacksmiths, weavers etc. Convergence of various schemes to attain this objective is also proposed."

The Ministry of Labour & Employment constituted Mentor Councils (MCs) to revamp courses run / to be run under National Council of Vocational Training (NCVT) in 25 sectors which will give a sustained skill based employability to the ITI trainees as the main objective of Vocational training. The ultimate approach of NIMI is to prepare the validated IMPs based on the exercises to be done during the course of study. As the skill development is progressive the theoretical content on a particular topic is limited to the requirement in every stage. Hence the reader will find a topic spread over a number of units. The test and assignment will enable the instructor to give assignments and evaluate the performance of a trainee. If a trainee possesses the same it helps the trainee to do assignment on his own and also to evaluate himself. The wall charts and transparencies are unique, as they not only help the instructor to effectively present a topic but also helps the trainees to grasp the technical topic quickly. The instructor guide enables the instructor to plan his schedule of instruction, plan the raw material requirement ,

Thus the availability of a complete Instructional Media Package in an institute helps the trainer and management to impart an effective training. Hence it is strongly recommended that the Training Institutes/Establishments should provide at least **one IMP** per unit. This will be small, one time investment but the benefits will be long lasting.

The **ICTSM trade Theory 1st semester in IT & ITES Sector** is one of the book develop by the core group members of the Mentor Councils (MCs). The 1st semester book includes **Module1-General & Electrical Safety, Module 2-Basic Electrical, Module 3-Basic Electronics,Module 4-Digital Electronics,Module 5-General & Electrical Safety.**

The **ICTSM Trade Theory 1st semester** is the outcome of the collective efforts of Members of Mentor Council which includes academic/professional institutions (IITs etc.), experts from field institutes of DGE&T, champion ITIs for each of the sectors, and also Media Development Committee (MDC) members and staff of NIMI.

NIMI wishes that the above material (Trade Practical & Trade Theory) will fulfil to satisfy the long needs of the Trainees and Instructor and helps the trainees for their employability in vocational training.

NIMI would like to take this opportunity to convey sincere thanks to all the Mentor Council members and Media Development Committee (MDC) members.

(A. MAHENDIRAN)
Director, NIMI.

ACKNOWLEDGEMENT

National Instructional Media Institute (NIMI) sincerely acknowledges with thanks for the co-operation and contribution extended by the following Media Developers and their sponsoring organisation to bring out this IMP **(Trade Theory)** for the trade of **ICTSM** under the **IT & ITES** Sector for Craftsman Training Scheme. This Book is prepared as per Revised Syllabus.

MEDIA DEVELOPMENT COMMITTEE MEMBERS

Dr. Sanjeev Kumar Gupta	-	Additional Director (IT), NIELIT, Deity, MCIT, New Delhi. Chairman, Mentor council.
Shri. Naresh Chandra	-	Joint Director of Training , DGE&T, New Delhi Mentor, Mentor council.
Dr. M. Jayaprakasan,	-	Deputy Director of Training, ATI, Chennai Team Leader, Mentor council.
Shri. St. Altaf Hossain	-	Training Officer, ATI, Howrah. Member, Mentor council.
Shri. P. T. Noushad	-	Assistant Training Officer, Govt. ITI (W), Coimbatore. Member, Mentor council.
Shri. Indu Jaisal	-	Senior Technical Leader MDC Member-NIMI. Member, Mentor council.
Shri. N. Sundararajan	-	Training Officer, NIMI, Chennai - 32 Co-ordinator, NIMI, Chennai.
Shri. V. Gopalakrishnan	-	Training Officer, NIMI, Chennai - 32 Co-ordinator, NIMI, Chennai.

NIMI records its appreciation of the Data Entry, CAD, DTP operators for their excellent and devoted services in the process of development of this instructional material.

NIMI also acknowledges with thanks, the invaluable efforts rendered by all other staff who have contributed for the development of this Instructional material.

NIMI is also grateful to all others who have directly or indirectly helped in developing this IMP.

INTRODUCTION

TRADE THEORY

The manual of trade theory consists of theoretical information for the First Semester course of the ICTSM Trade. The contents are sequenced according to the practical exercises contained in the manual on Trade practical. Attempt has been made to relate the theoretical aspects with the skill covered in each exercise to the extent possible. This co-relation is maintained to help the trainees to develop the perceptual capabilities for performing the skills.

The Trade Theory has to be taught and learnt along with the corresponding exercise contained in the manual on trade practical. The indications about the corresponding practical exercise are given in every sheet of this manual.

It will be preferable to teach/learn the trade theory connected to each exercise at least one class before performing the related skills in the lab. The trade theory is to be treated as an integrated part of each exercise.

The material is not for the purpose of self learning and should be considered as supplementary to class room instruction.

TRADE PRACTICAL

This manual of trade practical consists of a series of practical exercises and Lab assignments to be completed by the trainees during the first semester of the Information & Communication Technology System Maintenance (ICTSM) course. These exercises are designed to ensure that all the skills in the prescribed syllabus are covered. The manual is divided into Five modules. The distribution of time for the practical in the Five modules is given below.

Module 1	General & Electrical Safety	25 Hrs.
Module 2	Basic Electrical	160 Hrs
Module 3	Basic Electronics	180 Hrs
Module 4	Digital Electronics	100 Hrs
Module 5	Drives, Relays & Sensors	60 Hrs
Total		525 Hrs

While developing the practical manual, effort was made to prepare the exercises to be easily

understood and carried out by trainees whose level is below average. However the development team accepts that there is a scope for further improvement, NIMI looks forward to the suggestions from the experienced training faculty to improve the material.

CONTENTS

Lesson No.	Title of the Lesson	Page No.
	Module 1 : General & Electrical Safety	
1.1.01	Punctuality and Discipline	1
1.1.02	Artificial Respiration & Electrical safety	11
	Module 2 : Basic Electrical	
1.2.01	Basic electricity, fuses and switches	28
1.2.02	Metre range & measure types	35
1.2.03	Simple AC, DC Circuits and measure current & voltage	38
1.2.04	Multimeter	40
1.2.05	Skinning and tinning of wires	45
1.2.06	Soldering & De - Soldering Technique	49
1.2.07	Resistors and finding its value	58
1.2.08	Ohm's law	68
1.2.09	Kirchhoff's Laws	74
1.2.10	Electrical power	78
1.2.11	Special types of resistors	83
1.2.12	Working principle of oscilloscope	91
1.2.13	Inductors and Inductive Reactance at different frequencies	98
1.2.14	Concept of Self & Mutual Inductance	109
1.2.15 - 1.2.16	Transformers and its Types	112
1.2.17	Voltage regulation, efficiency and phase relationship in transformers	123
1.2.18	Capacitors - construction & testing	127
1.2.19	Measurement of capacitive Reactance at different frequencies	138
1.2.20	Capacitors in series & parallel	144
1.2.21	Types of cells & Batteries	156
1.2.22	Series & parallel connection of batteries	163
1.2.23	Charging techniques of secondary batteries	168

Lesson No.	Title of the Lesson	Page No.
	Module 3 : Basic Electronics	
1.3.01	Characteristics & specification of Diodes	175
1.3.02	Half wave & Full wave Rectifiers	184
1.3.03	Bridge Rectifiers	188
1.3.04	Filters & Rectifiers	191
1.3.05	Characteristics & specification of zener diodes	200
1.3.06	Zener diodes - Designing regulators	203
1.3.07	Identify the test different types transistors	205
1.3.08	Biasing of Transistors and common Emitter characteristics	212
1.3.09	Gain and Impedance of common Emitter amplifier	218
1.3.10	Emitter follower / Common Collector	227
1.3.11	Common base amplifier	230
1.3.12	Field effect transistor and its types	232
1.3.13	Unijunction Transistor (UJT) its applications	238
1.3.14	Working of SCR, TRIAC and DIAC	241
1.3.15	Crowbar protection using SCR(s)	249
1.3.16	AC power control using TRIAC	251
1.3.17	Series voltage regulators	255
1.3.18	Fixed & variable regulators using 3 pin ICs	261
1.3.19	Working of inverters	270
1.3.20	Working principle of oscilloscopes	276
1.3.21	Measurement using dual trace oscilloscope - 1	282
1.3.22	Controls and functions in oscilloscopes	283
1.3.23	Simple AM Transmitter	288
1.3.24	Crystal AM receiver	292
1.3.25	Frequency modulation and FM receivers	298

Lesson No.	Title of the Lesson	Page No.
	Module 4 : Digital Electronics	
1.4.01	Digital IC families and their operational characteristics	302
1.4.02	Number systems and conversions	309
1.4.03	OR and NOR gate circuits	313
1.4.04	AND, NAND and XOR gate circuits	315
1.4.05	R.S Flip-flops and D Flip-flop using NAND gate	318
1.4.06	J.K Flip-flop	327
1.4.07	Up/down counters	332
1.4.08 - 1.4.09	Binary arithmetic	335
1.4.10	Multiplexers & Demultiplexers	341
1.4.11	Analog-to-digital converter	343
1.4.12	Interfacing the digital and analog signals	345
1.4.13	Shift registers, Types, applications	349
	Module 5 : Drives, Relays & Sensors	
1.5.01	Identify and use of Gears, Belt, Stepper Motor, Drive.	353
1.5.02	Electrical / Electronic & Mechanical Sensors	357
1.5.03	Advanced Intel microprocessor chips	365

Punctuality and Discipline

Objectives : At the end of this lesson you shall be able to

- **importance of punctuality**
 - **punctuality at work place**
 - **importance of discipline.**
-

Punctuality is the characteristic of being able to complete a required task before or at a previously designated time

Importance of punctuality

Being punctual strengthens and reveals your integrity. If you tell someone that you will meet them at a certain time, you have essentially made them a promise. And if you say you'll be there at 8:00, and yet arrive at 8:15, you have essentially broken that promise. Being on time shows others that you are a man of your word.

Being punctual shows you are dependable. A man can always be found at his post, carrying out the duties needful for that time. People know they can rely on such a man – if he says he will be there, he'll be there. But if a man is not punctual, others cannot depend on him — they do not know where he will be when they need him. His associates will begin to feel he cannot organize his own time, and these doubts will seep into matters beyond the clock, as it naturally raises the question: "If he is careless about time, what else is he careless about?"

Benjamin Franklin once said to an employee who was always late, but always ready with an excuse: "I have generally found that the man who is good at an excuse is good for nothing else."

Being punctual builds your self-confidence. Showing up on time not only tells other people you are dependable, it teaches you that you can depend on yourself. The more you keep the promises you make, the more your self-confidence will grow. And the more you gain in self-mastery, the less you will be at the mercy of your compulsions and habits, and the more in control of your life you will feel.

Being punctual assures you're at your best. After riding someone's bumper, speeding like a maniac, scanning for cops, and cursing at red lights, it's hard to then turn your focus to concentrate in the classroom. But when you show up on time, better yet a little early, you have a few minutes to collect your thoughts, and review your materials

Being punctual builds and reveals your discipline. The punctual man shows that he can organize his time, that he pays attention to details, and that he can put aside this to do that – he can set aside a pleasure to take care of business.

Being punctual shows your respect for others. Being late is a selfish act, for it puts your needs above another's. You want an extra minute to do what you'd like, but in gaining that minute for yourself, you take a minute from another, which is why....

Being late is a form of stealing. That's a tough truth, but it's a truth nonetheless. When you make others wait for you, you rob minutes from them that they'll never get back. Time they could have turned into money, or simply used for the things important to them. In coming to meet you at the agreed upon hour, they may have made sacrifices – woken up early, cut short their workout, told their kid they couldn't read a story together – and your lateness negates those sacrifices. If you wouldn't think of taking ten from another man's wallet, you shouldn't think of stealing ten minutes from him either. Being punctual shows you value time yourself, and thus wouldn't think of depriving others of this precious, but limited resource.

Being late disturbs the experiences of other people. Your tardiness not only robs others of their time, but of the fullness of their experiences as well. The student who interrupts in the middle of his demonstration the family which climbs over you to get to their seats at the middle of the row in the theater; the man who opens the creaky door in the middle of a eulogy. When an old man was once asked why he had been so punctual in arriving at his church on time for decades, he replied, "I made it my religion not to disturb the religion of others".

Being late strains your relationships. When you're late in meeting other people, it makes them feel under-valued, that whatever you couldn't pull yourself away from was more important or that they didn't mean enough to you to warrant allotting sufficient time to arrive on schedule. The guest who flies in to see you feels like a dope standing at the airport alone, your date feels awkward sitting at the restaurant by herself, and your child feels abandoned as she waits with her teacher for you to arrive, all the other children having already been picked up from school.

Being late hurts your professional career. Whether you're an employee or in business for yourself, being late can hinder your professional success. Many companies have strict policies about punctuality — get a few write-ups

and you're gone. Of course, if you arrive late to the job interview, you probably won't land the position in the first place. And if you're trying to win over a new client, arriving ten minutes late isn't going to get things off on the right foot, in the same way that promising to get something to him by a certain date and then failing to do so, may have him looking elsewhere for your services.

Being late takes a toll on your life. Always running behind simply hurts you in all areas of your life. It results in lost opportunities: missing a plane, missing a meeting, missing an important demonstration class, missing a wedding. It results in embarrassment and forces you to come up with excuses for why you're late, putting a strain on your honesty. Basically, it makes your life more complicated; for men seeking to simplify their lives, cultivating punctuality is an essential part of that path.

punctuality at work place

- 1 Entering the lab / workshop in time
- 2 Having every thing ready before starting the task/ practicals
 - plan the sequence of work
 - layout the tools yours, in the order of sequence
- 3 Complete the task within the allotted time

Importance of discipline

Discipline is essential for every group, for every society and for every political institution. Without discipline there can be no society or Government. No nation can exist without discipline. It is discipline that unites man to man, and one society to another. Hence discipline is one of the basic insignia of social life.

Discipline has to be observed in speech, in sport and in every kind of relationship. For instance, during bhajans, all those singing in chorus have to maintain the same tune. Discordant singing will jar on the ears. Singing in unison in bhajans is a form of discipline.

In games, you have an umpire to enforce the rules of the game. Every player has to observe the rules strictly. Sometimes while playing, in their enthusiasm, the players fail to observe the rules. The umpire, however, sees to it that the rules are enforced and the players have to obey him implicitly. Any player who does not obey will be violating the rules of the game.

Discipline means the observance of certain well-defined rules. Without such regulation it is not possible to maintain humanness. Such regulation contributes to the glory of human existence. Discipline cannot be acquired from books. Nor can it be learnt from teachers. It has to be as a natural component of one's daily life in the discharge of one's duties. Discipline is essential from the moment of waking to the time of going to sleep.

Today it is because people do not observe the rules, whether in sports or elsewhere, life has become intolerable. Some persons hold positions of authority in

business, in administration or in other institutions. Here also discipline is essential, but because one holds an office, he cannot behave as he likes; he cannot be free with his tongue. In his words and his writings he should observe restraint. In the use of words, whether in speech or writing, care should be observed regarding their future implications. Your discipline will protect you in whatever you do. Even in a simple matter like walking on the road, one may do not observe the rules. They choose to walk on the middle of the road instead of using the pavement, jaywalking on the road is not only harmful to oneself but is a hindrance to others. Your motto in life should be: Help over, hurt never.

Discipline must start from the early years, but it is needed not only by students but by elders and all persons in authority.

Discipline and punctuality are two most essential traits required in a professional to be successful. Discipline ensures individuals behave in an acceptable way at the workplace and also adhere to the rules and regulations of the organization. Individuals who are disciplined are not only successful professionally but also in their personal lives. Disciplined employees are liked and appreciated by not only their superiors but also other fellow workers. Such employees climb the success ladder quickly as compared to individuals who attend office just for the sake of it and they often find themselves out of the system in no time. Misbehaving at the workplace spoils the entire work culture eventually leading to zero or very less productivity. You need to understand that there is a difference between office and home and you just can't behave in the same way at work as you behave at home. Discipline is crucial at workplace as it ensures employees behave in a decent way and also maintain the decorum of the workplace.

Examples of indiscipline:

- 1 Excessively late, slow or sluggish
- 2 Failure to notify of an absence
- 3 Insubordination willfully disobeying the superior
- 4 Using rude or abusive language
- 5 Failure to follow rules procedures or policies
- 6 Dishonesty
- 7 Theft of tools, equipments consumables etc

Safety in moving and shifting heavy and delicate equipments

Objectives: At the end of this lesson you shall be able to

- state the types of injury caused while lifting and carrying loads, and how to prevent them
- state the six points to be noted in the kinetic method of lifting.

Many of the accidents reported involve injuries caused while lifting and carrying loads. An electrician may need to install a heavy electric motor in a restricted space and make the correct wiring connections. Wrong lifting techniques can result in injury. A load need not necessarily be very heavy to cause injury.

What causes an injury more frequently -
a very heavy load or a wrong way of lifting?

A wrong way of lifting something too heavy causes injury. The muscles and joints are strained. This is especially true in the case of injury to the back. The most common type of injury due to a wrong way of lifting is injury to the back.

Are the injuries caused as a result of strain only?

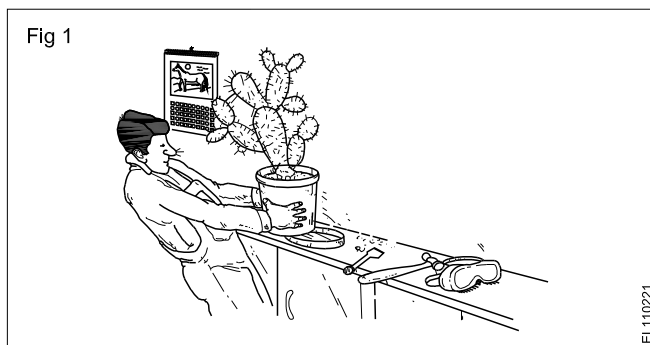
No. Injuries during lifting and carrying may be caused by tripping over an object and falling, or striking an object with a load.

Types of injury and how to prevent them

Cuts and abrasions: Cuts and abrasions are caused:

- by rough surfaces and jagged edges
- by splinters and sharp or pointed projections.(Fig 1)

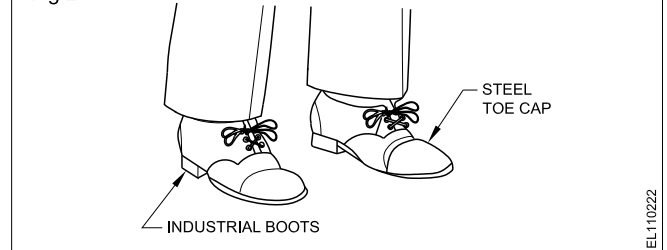
Leather gloves will usually be sufficient for protection, but the load should be checked to make sure of this, since large or heavy loads may involve body contact as well.



Crushing of feet or hands: Feet or hands should be positioned so that they cannot be trapped by the load. Timber wedges can be used when raising and lowering heavy loads to ensure that the fingers and hands are not caught and crushed.

Safety shoes with steel toe caps will protect the feet.(Fig2)

Fig 2

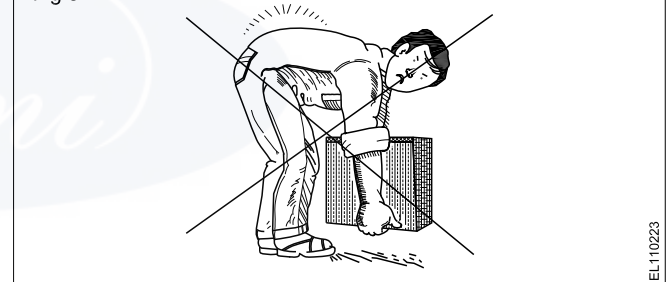


Strains to muscles and joints: Strains to muscles and joints may be the result of:

- lifting a load which is too heavy or
- lifting incorrectly.

Sudden and awkward movements such as twisting or jerking during a lift can put severe strain on muscles. Injury to the back is most often caused by lifting loads which are too heavy or by employing wrong methods. 'Stoop lifting' - lifting from a standing position with the back rounded (Fig 3) increases the risk of a back injury.

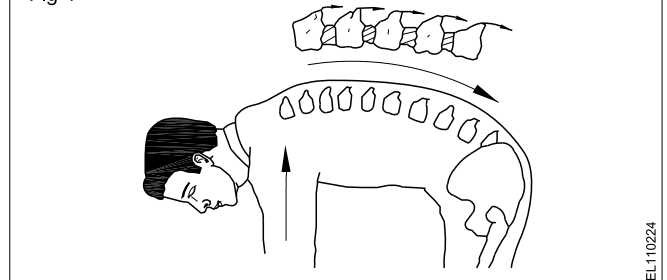
Fig 3



The human spine is not an efficient weight lifting machine, and can be easily damaged if incorrect techniques are used.

The stress on a rounded back can be up to six times greater than if the spine is kept straight. Fig 4 shows an example of stoop-lifting.

Fig 4



Preparing to lift: What preparation should you have to make before lifting?

Before lifting or handling any load ask yourself the following questions.

What has to be moved?

Where from and where to?

Will assistance be required?

A load which seems light enough to carry at first will become progressively heavier, the farther you have to carry it.

Before lifting and carrying, make sure the route is clear of obstacles and the place where the load is to be deposited is not obstructed.

The person carrying the load should always be able to see over or around it.

The weight, a person can lift will vary according to:

- age
- physique and
- condition such as other health factors.

It will also depend on whether one is used to lifting and handling heavy loads.

What makes an object difficult to lift and carry?

- Weight is not the only factor which makes it difficult to lift and carry.
- The size and shape can make an object awkward to handle.
- Loads, which require the arms to be extended in front of the body, place more strain on the back and stomach than do compact objects carried close to the body.
- The absence of hand-holds or natural handling points can make it difficult to raise and carry the object.

The kinetic method of lifting: It enables the worker to make full use of the body's own weight to initiate the lift.

The natural shape of the spine is maintained throughout (although the body may be bent forward, the spine should remain straight), and the lift is powered by the strong muscles in the legs and thighs.

It is important to begin with the right posture, that is, the various parts of the body should be correctly positioned before beginning the lift. The following six points should be noted.

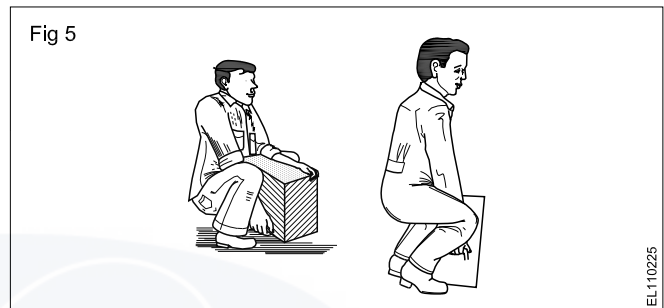
- The Feet are placed about 18 inches(45cms) wide apart, with one foot slightly forward, in the direction of movement. This gives a good balance and provides a secure basis for the lift.
- The Knees should be slightly bent (but not fully bent as in a squat).

- The Back must be straight, although the body may be inclined forward as shown in the illustration.
- The Arms should be as close to the body as possible. The farther the arms are extended the greater the strain. Elbows too should be kept in.
- The Grip must be firm and secure.
- The Head should be erect with the chin in.

Correct manual lifting techniques

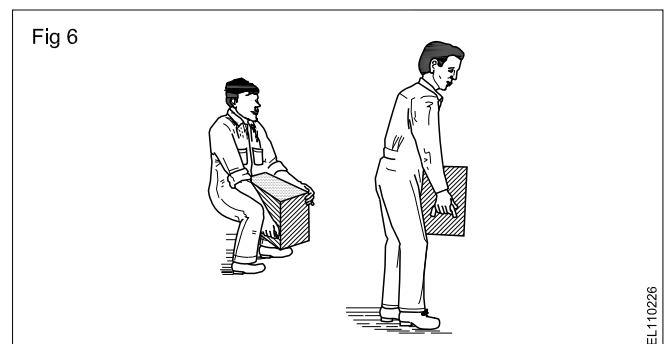
- Approach the load squarely, facing the direction of travel.
- The lift should start with the lifter in a balanced squatting position, with the legs slightly apart and the load to be lifted held close to the body. Ensure that a safe secure hand grip is obtained. Before the weight is taken, the back should be straightened and held as near the vertical position as possible. (Fig 5)

Fig 5



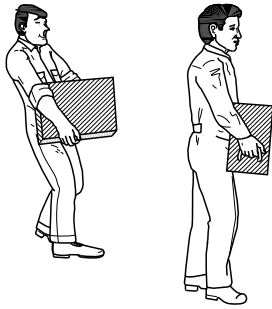
- To raise the load, first straighten the legs. This ensures that the lifting strain is being correctly transmitted and is being taken by the powerful thigh muscles and bones. Look directly ahead, not down at the load while straightening up, and keep the back straight; this will ensure a smooth, natural movement without jerking or straining. (Fig 6)

Fig 6



- To complete the lift, raise the upper part of the body to the vertical position. When a load is near to an individual's maximum lifting capacity, it will be necessary to lean back on the hips slightly (to counterbalance the load) before straightening up. (Fig 7)

Fig 7



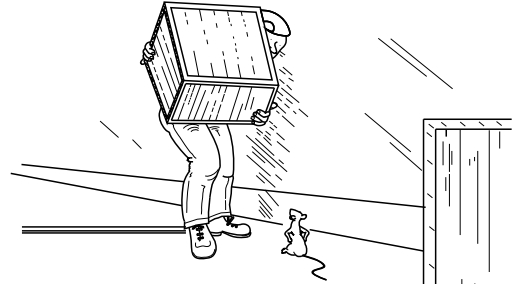
EL110227

- Keeping the load well into the body, carry it to the place where it is to be set down. When turning, avoid twisting from the waist—turn the whole body in one movement.

Lowering the load: Make sure the area is clear of obstructions. (Fig 8)

Bend at the knees to a semi-squatting position; keep the back and head erect by looking straight ahead, not down at the load. It may be helpful to rest the elbows on the thighs during the final stage of lowering. (Fig 7)

Fig 8



EL110228

Moving heavy equipment

Objectives: At the end of this lesson you shall be able to

- name the methods followed in industry to move heavy equipment
- describe the procedure to be followed for moving heavy equipment on layers and rollers
- list the safety considerations while raising a load and moving a load.

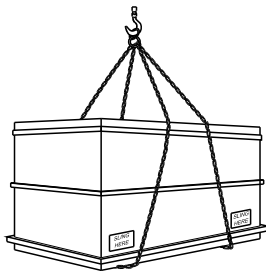
Heavy equipment is moved in industry using any of the following

- Cranes and slings
- Winches
- Machine moving platforms
- Layers and rollers.

Cranes and slings: This method is used whenever loads are to be lifted and moved. (Fig 1) Examine the slings for cuts, abrasion, wear, fraying or corrosion.

Damaged slings must not be used.

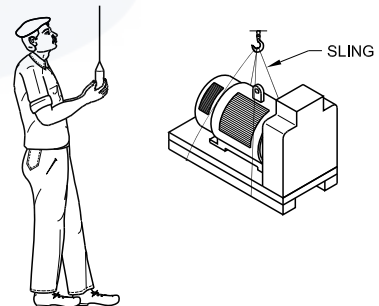
Fig 1



EL110231

Distribute the weight as evenly as possible between the slings when using more than one sling. (Fig 2)

Fig 2



EL110232

Keep the slings as near to vertical as possible.

Winches: Winches are used to pull heavy loads along the ground.

Cranes and slings may be power-driven (Fig 3) or hand-operated. (Fig 4)

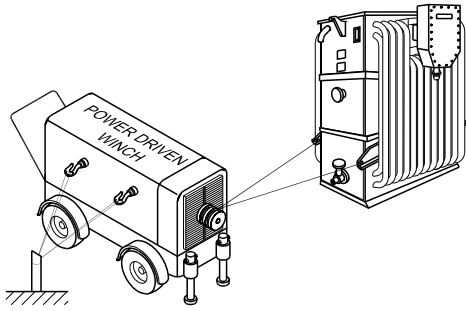
Ensure that the safe working load (SWL) of the winch is adequate for the task.

Secure the winch to a structure which is strong enough to withstand the pull.

On open ground, drive long stakes into the ground and secure the winch to them.

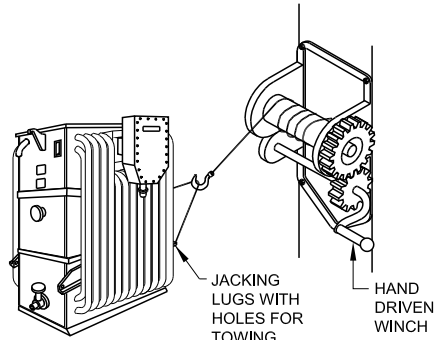
Choose a suitable sling and pass it around the base of the load. Secure it to the hook of the winch.

Fig 3



EL110233

Fig 4



EL110234

Some heavy items have special lugs welded to them for jacking and towing purposes.

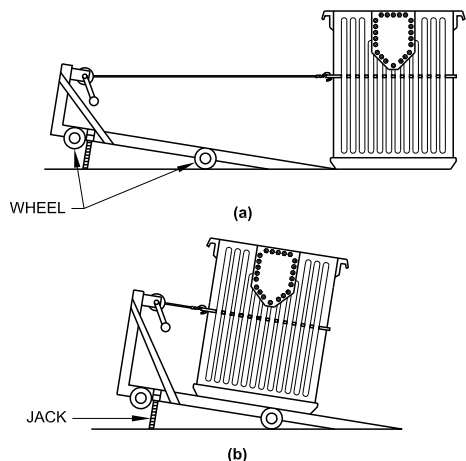
Safety considerations: Before using any winch, check that the brake or ratchet mechanism is in working order. You should know how to control it.

Keep hands and fingers well away from the gear wheels.

Keep bearings and gears oiled or greased.

Machine moving platforms: This is a special device made to move heavy equipment in industry. Figs 5a and 5b show the method of loading the equipment.

Fig 5



EL110235

Pass a suitable sling round the load at a convenient height.

Attach the sling to the hook of the winch and draw the load on the platform until its centre of gravity lies between the front and rear wheels.

Lower the jacks so that the platform rests on its wheels.

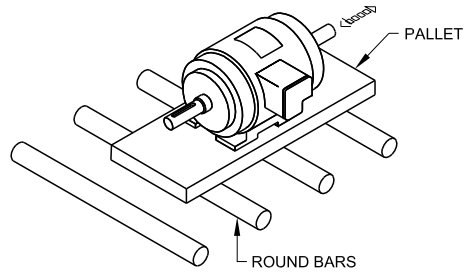
For unloading follow the procedure in the reverse order.

Using layers and rollers: Sometimes a load cannot be moved along the ground because of the irregular shape of its base or because it is not rigid enough.

Place such a load on a flat-bottomed pallet or 'layer' resting on round bars. (Fig 6)

Ensure the bars (rollers) are long enough to project at each side of the load, for ease of handling.

Fig 6

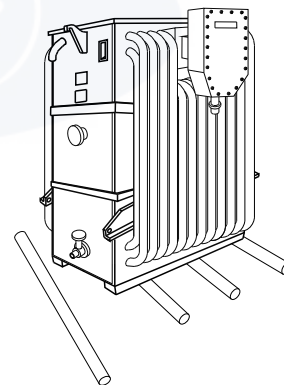


EL110236

They should be large enough to roll easily over any unevenness along the route but should be small enough to be handled easily. (Fig 7)

Two or three bars of equal diameter are sufficient for most loads but if four or more bars are used the load may be moved faster as there is no delay when moving the rear bar to the front. (Fig 7)

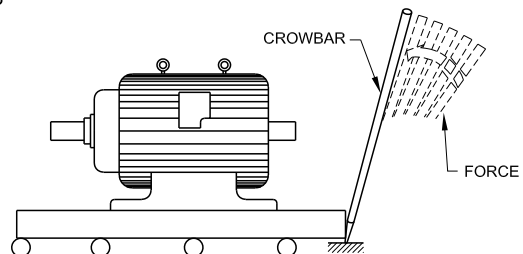
Fig 7



EL110237

Move the load by using a crowbar. (Fig 8)

Fig 8



EL110238

Caution:

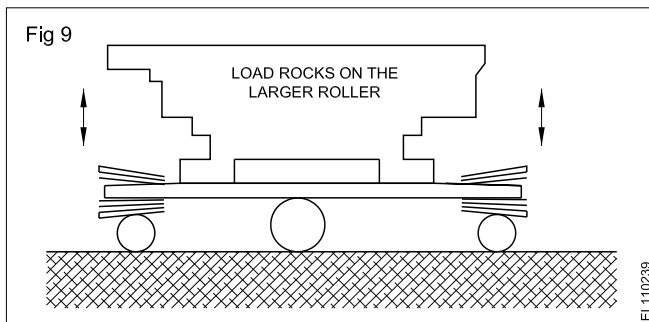
When a load is on rollers, only shallow slopes can be negotiated.

Hold the load in check all the time it is on the slope.

Use a winch with an effective brake for this operation.

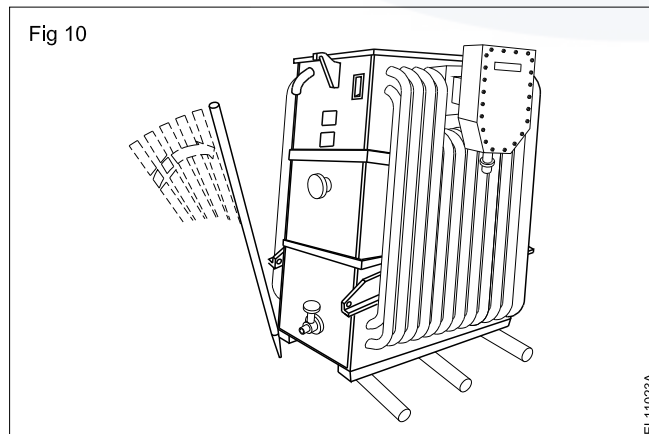
To negotiate a corner on rollers: For a moderate load, insert one roller a little larger than the others as the corner is approached.

When this roller is under the centre of gravity of the load, the load can be rocked to and fro on the roller and swivelled around sideways. (Fig 9)



For heavier loads: Stop the load at the beginning of the corner.

Twist the load round on the rollers by pushing the sides with a crowbar until the load is just over the ends of the rollers. (Fig 10)



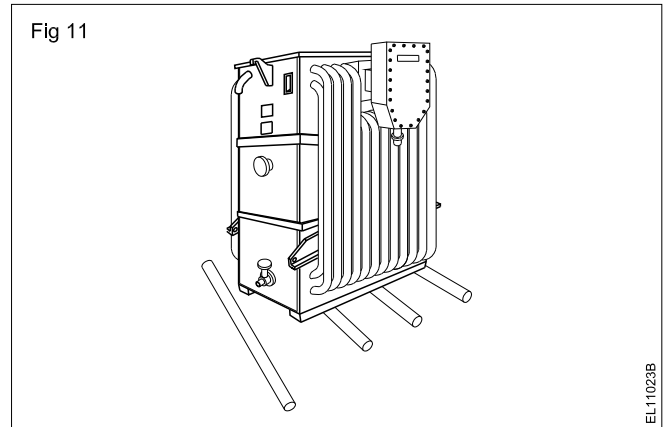
Place some rollers at an angle to the front of the load.

Push the load forward on to these rollers.

Twist the load further round and place the freed rollers in front of and at an angle to the load. (Fig 11)

Continue until the load is pointing in the desired direction.

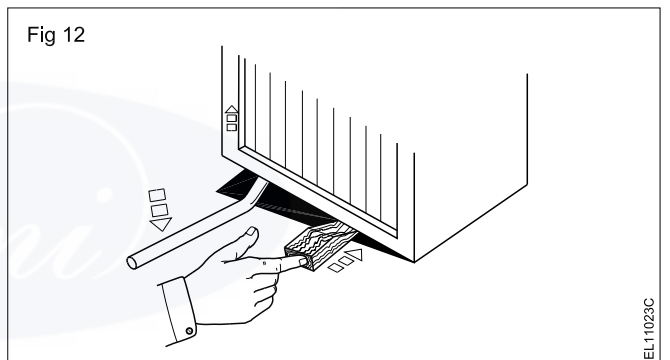
Fig 11

**Safety considerations**

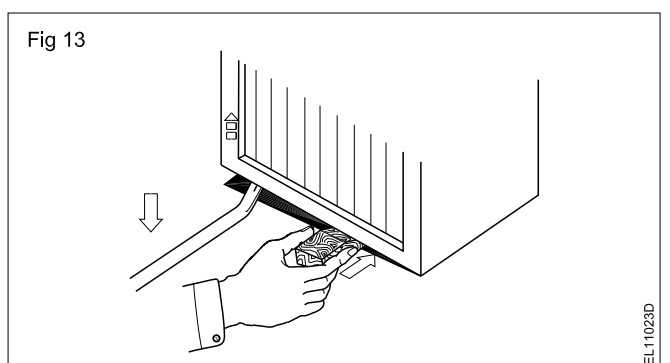
Moving heavy loads with crowbars or jacks: Make sure hands are clear of the load before lowering it on to packing or rollers.

Do not use hands underneath the packing when positioning it. Use a push block.

Place the packing on the floor and push it under the load. (Fig 12)



Hold it by its side faces keeping the fingers well away from the lower edge of the load and from the floor. (Fig 13)



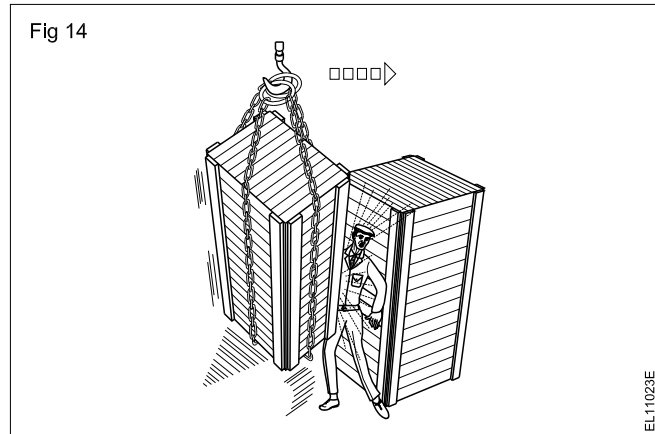
Raising a load: Check that the slings are correctly secured to the load and to the hook. Ensure they are not twisted or caught on a projecting part of the load.

If you cannot see an assistant on the far side of the load, verify that he is ready before starting to lift and ensure that his hands are clear of the slings.

Warn nearby workers that the lift is about to begin.

Lift slowly !

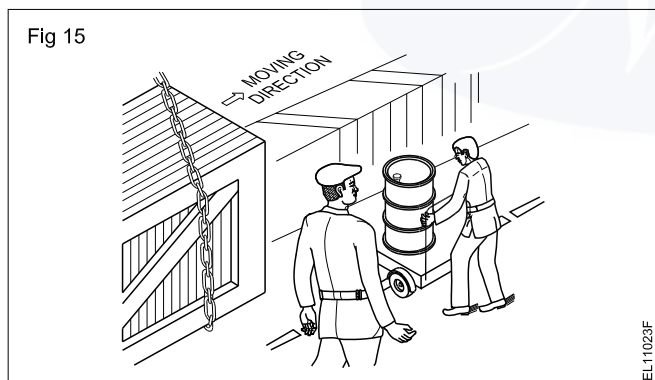
Take care to avoid being crushed against other objects as the load rises. It may swing or rotate as it leaves the ground. (Fig 14)



Minimise such movement by locating the hooks as accurately as possible above the centre of gravity of the load. Keep the floor clear of unnecessary objects.

Moving a load: Check that there are no obstacles in the way of the crane and the load. (Fig 15)

Stand clear of the load and move it steadily.



Safety signs

Objectives : At the end of this lesson you shall be able to

- state the common causes for accidents
- state the functions of signboards
- name the four basic categories of safety signboards
- give examples of each category of signboards.

Causes for accidents

Accidents happen due to causes which can not be foreseen. Most accidents are a result of human error, ignorance, neglect, forgetfulness or recklessness. Most accidents can be prevented.

Normally accidents do not just happen. They are caused.

Causes for accidents are many, some of the important ones are:

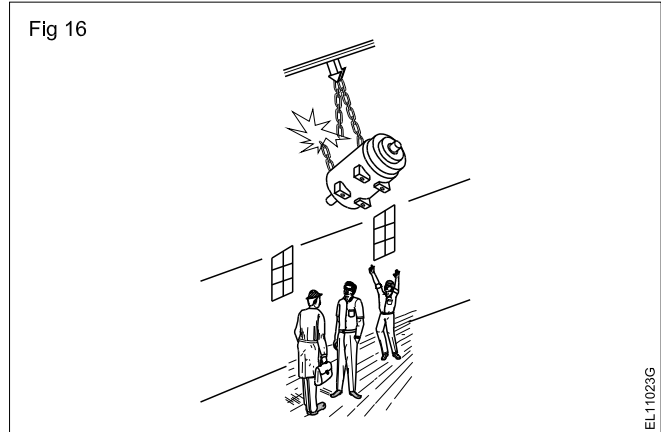
- unawareness of danger
- disregard for safety
- negligence
- lack of understanding of proper safety procedures
- untidy condition of workplace
- inadequate light and ventilation
- improper use of tools

Be prepared to stop the load quickly, if somebody moves into its path.

Allow for the natural swing of the load when changing speed or direction.

Ensure that the load will not pass over the heads of other people. (Fig 16)

The tackle may fail or slip.



Warn them to stand clear before an accident is caused.

Accidents do not happen. They are caused.

Safe attitudes

People's attitudes govern what **they do** or **fail to do**. In most cases accidents occur when someone is working with unsafe equipment or in an unsafe situation. These situations which cause accidents is because the working person has failed to do certain things.

Most accidents do not just happen. Accidents are caused by those who, for example, damage an equipment or find that it is faulty but don't set it right or atleast report it to the concerned people. Anybody who sees a hazard and does nothing about it is contributing to the possibility of an accident. Knowing the possibility of an accident , if a worker does not bother about it, that by itself is enough to ensure a workman is crippled for life. He might not have caused the accident, but could have definitely prevented it.

Safety signboards

Signboards are a common sight in almost all places such as roadways, railways, hospitals, offices, institutions, industrial units and so on.

Signboards are visual indicators. The signs on the signboards may be just a symbol, a small text, a figure or a combination of these.

Signboards carry a single clear message. These messages are to ensure safety.

Signboards can be classified into four basic categories.

Prohibition signs

Indicating a behaviour which is prohibited (not allowed) in that situation or environment. Refer to Chart 1 for examples.

Mandatory signs

Indicating a behaviour which is a must, which when not obeyed may cause accidents. Refer to Chart 1 for examples.

Warning signs

Indicating a warning such that suitable precaution is taken. Refer to Chart 1 for examples.

Information signs

Giving information which is very useful and reduces waste of time. Refer to Chart 1 for examples.



CHART 1 SIGN BOARDS

a) PROHIBITION SIGNS



Shape	Circular.
Colour	Red border and cross-bar. Black symbol on white background.
Meaning	Shows what <i>not</i> be done.
Example	No smoking.

b) MANDATORY SIGNS

Shape	Circular.
Colour	White symbol on blue background.
Meaning	Shows what <i>must</i> be done.
Example	Wear hand protection.

c) WARNING SIGNS

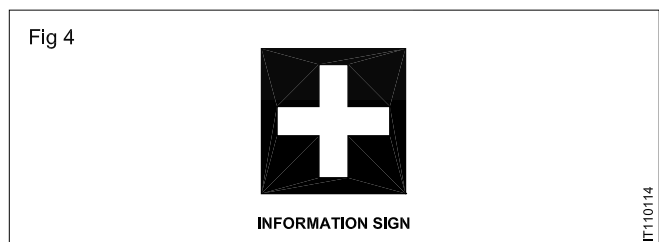


Fig 2



Shape	Triangular.
Colour	Yellow background with black border and symbol.
Meaning	Warns of hazard or danger.
Example	Caution, risk of electric shock.

d) INFORMATION SIGNS



Shape	Square or oblong
Colour	White symbols on Green background.
Meaning	Indicates or gives information of safety provision/First aid
Example	First aid point.

Artificial Respiration & Electrical safety

Objectives: At the end of this lesson you shall be able to

- state the importance of respiration
- state occasions when artificial respiration becomes a necessity
- state the various methods of artificial respiration
- list the procedure to be followed in different methods of artificial respiration
- state the action to be taken to treat a victim with cardiac arrest
- state the important action to be taken when a patient recovers after undergoing artificial respiration
- state the meaning of electric shock
- state the factors that determine the severity of a shock
- state the severity of shock based on the voltage level.

RESPIRATION

Respiration is an involuntary function of all living organisms. In human beings respiration involves breathing in and breathing out air. The air taken consists of oxygen which is made use of by the lungs to purify the blood. If a person cannot breathe or stops breathing for any reason he will collapse in no time.

Artificial respiration

There are occasions when breathing/respiration becomes difficult or feeble in a person due to sickness, shock or accident. When breathing becomes feeble the person becomes unconscious. At this stage it is possible to improve the respiration by artificial means. This is known as artificial respiration.

In artificial respiration, air containing oxygen is forced into the lungs of the unconscious person so as to maintain the supply of oxygen to the body till the time the victim reaches the hospital/clinic for further treatment.

There are several methods by which artificial respiration can be given. Some of the well known methods are,

- HOLGEN-NELSON method
- SCHAFER'S method
- Mouth-to-mouth method.

Basically all these methods serve the same purpose and are equally good. Therefore the person/volunteer can adopt any one of the above methods in which he is trained and is confident. However, Nelson's method and Schafer's method should be avoided while giving artificial respiration to those patients who are suspected to have injuries in the chest wall and abdomen.

Each of the above said methods has a systematic procedure to be followed. These procedures are given below. A person must practise this procedure thoroughly before giving artificial respiration to a sick person. Any short cuts in the procedure is dangerous and may worsen the condition of the sick person instead of improving it.

NELSON's Method (Arm-lift back-pressure method)

PRECAUTION

Nelson's arm-lift back pressure method must not be used in case of suspected injuries to the chest wall or abdomen of the victim.

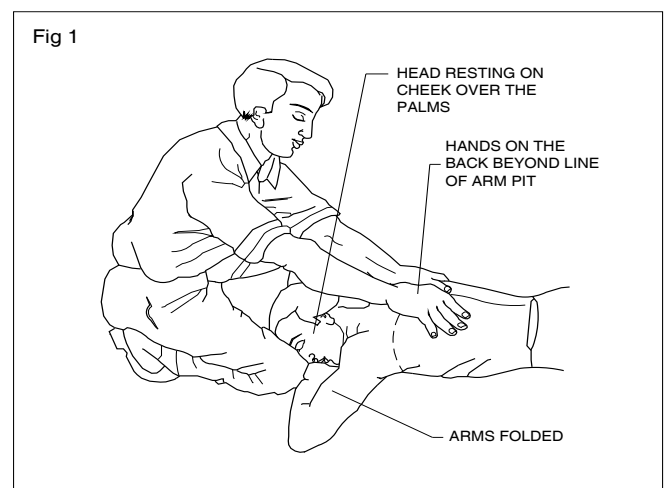
Be brisk in carrying out this method but avoid violent operations which may cause injury to the internal parts of the victim.

- 1 Loosen the clothing of the victim as tight clothing interferes with the victim's breathing.
- 2 Remove any foreign materials or false teeth from the victim's mouth, and keep the mouth open.

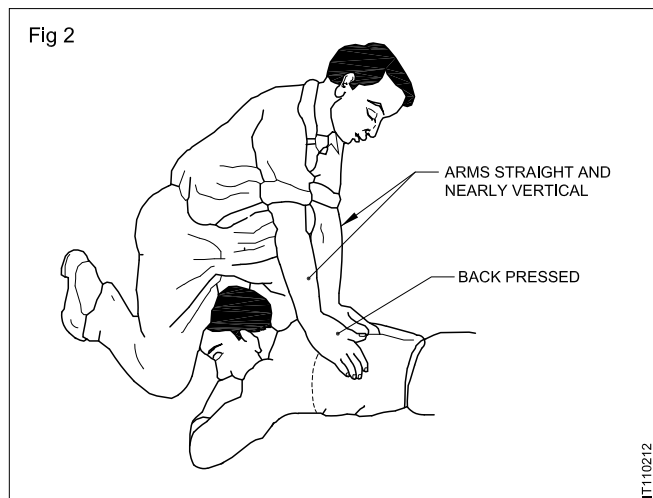
Do not delay artificial respiration for loosening clothes or even if the mouth is closed tightly.

- 3 As shown in Fig 1, place the victim face down with his arms folded, palms one over the other and head resting on his cheek over the palms. Kneel on one or both knees at the victim's hand. Place your hands on the victim's back beyond the line of the armpits. Spread your fingers outwards and downwards with the thumbs just touching each other.

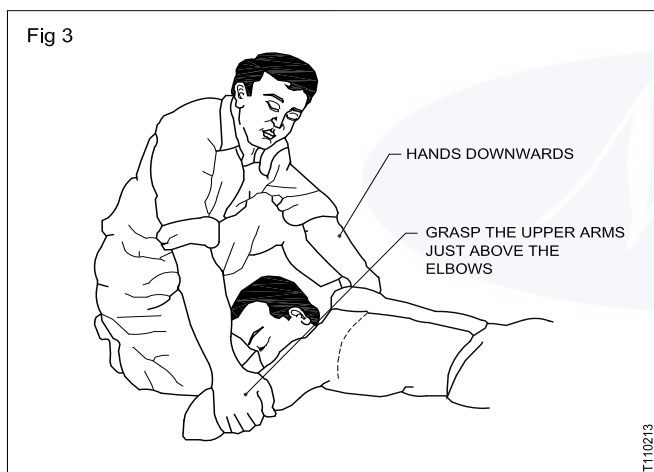
Fig 1



- 4 As shown in Fig 2, gently rock forward the arms keeping them straight until they are nearly vertical, and thus steadily pressing the victim's back as shown to force the air out of the victim's lungs.



- 5 As shown in Fig 3, synchronizing the above movement rock backwards sliding your hands downwards along the victim's arms. Grasp his upper arm just above the elbows as shown.



- 6 Now rock backwards. As you rock backwards, gently raise and pull the victim's arms towards you as shown in Fig 4 until you feel tension in his shoulders. Remain in this position for a few seconds. To complete the cycle, lower the victim's arms and move your hands up to the initial position.
- 7 Repeat the cycles a few more times by following steps 3 to 6.

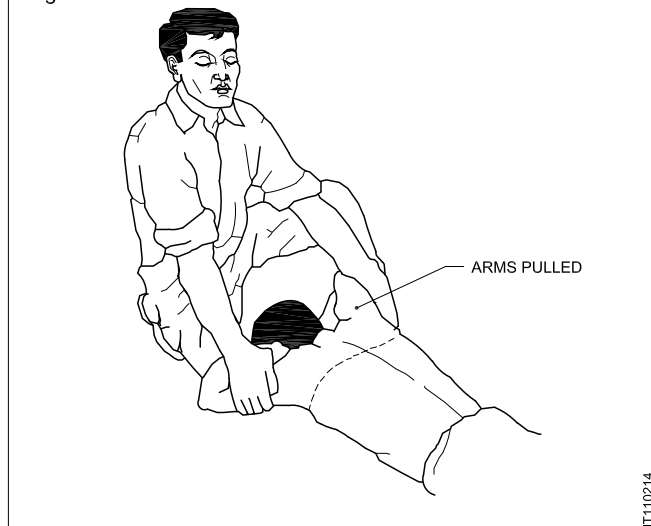
SCHAFER'S method

PRECAUTION

Do not use this method of artificial respiration in case the victim has injuries on his chest or abdomen.

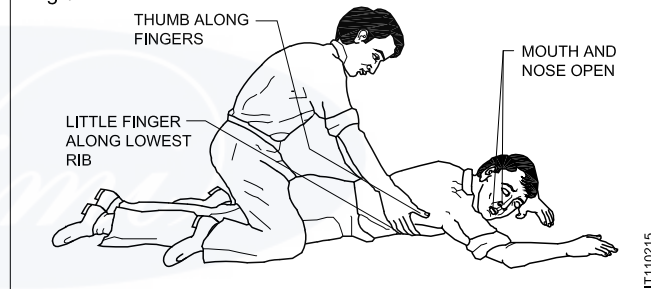
Be brisk in carrying out this method but avoid violent operations which may cause injury to the internal parts of the victim.

Fig 4



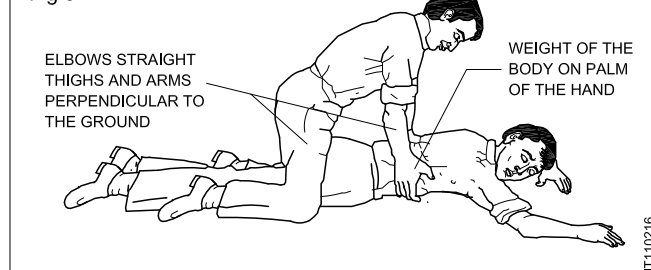
- 1 Loosen the clothing of the victim as tight clothing interferes with the victim's breathing.
- 2 Lay the victim on his abdomen. Extend one arm directly forward, and the other arm bent at the elbow as shown in Fig 5. Keep the face turned sideways and resting on the hand or forearm as shown in Fig 5.

Fig 5

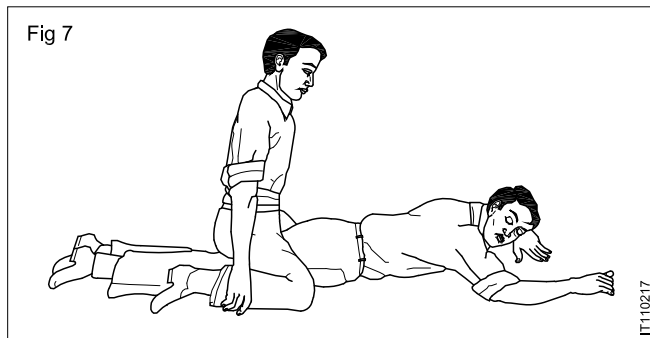


- 3 Kneel astride the victim as shown in Fig 5 such that his thighs are between your knees. Position your fingers and thumb as shown in Fig 5.
- 4 With the arms held straight, swing forward slowly so that the weight of your body is gradually applied on the lower ribs of the victim as shown in Fig 6. This weight force the air out of the victim's lungs.

Fig 6



- 5 Now swing backward immediately removing all pressure on the lower ribs as shown in Fig 7. This allows the lungs to get filled with air.



- 6 After two to three seconds, swing forward again and repeat the cycle (steps 4 and 5) twelve to fifteen times a minute.

Mouth-to-mouth method of artificial respiration

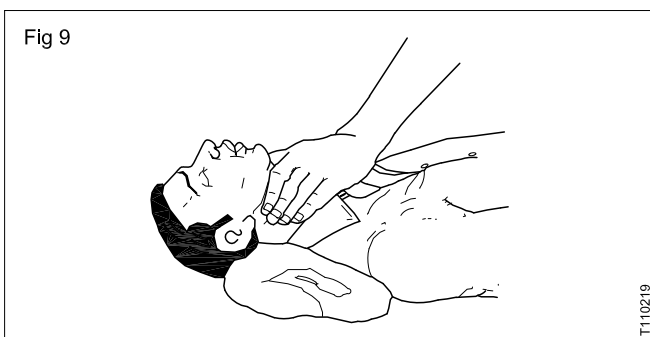
PRECAUTION

Be brisk in carrying out this method but avoid violent operations which may cause injury to the internal parts of the victim.

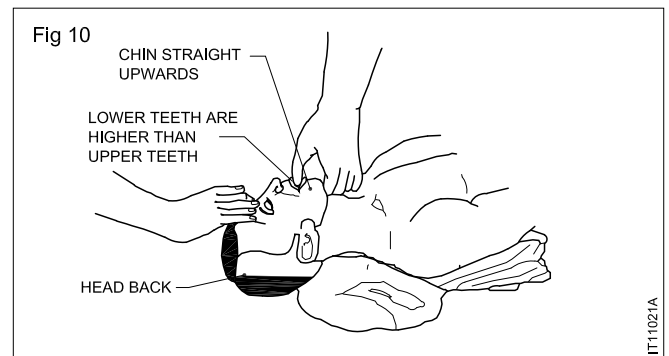
- 1 Loosen the clothing of the victim as tight clothing interferes with the victim's breathing.
- 2 Remove loose dentures or other obstructions from the mouth. Make sure that the victim's nose and mouth are clear.
- 3 Lay the victim flat on his back. Place a roll of clothing under his shoulders such that his head is thrown well back as shown in Fig 8.



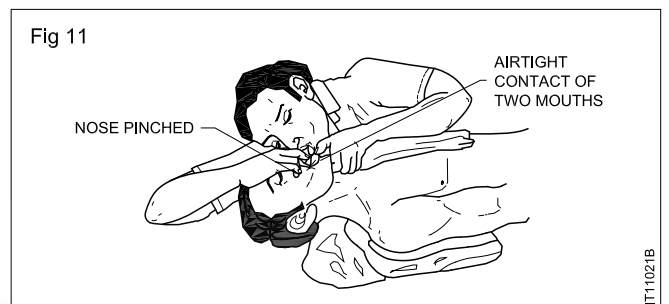
- 4 Tilt the victim's head back so that the chin points straight upward as shown in Fig 9.



- 5 Grasp the victim's jaw as shown in Fig 10, and raise it upward until the lower teeth are higher than the upper teeth. Maintain this position throughout the artificial respiration to prevent the tongue from blocking the air passage.



- 6 Take a deep breath and place your mouth over the victim's mouth as shown in Fig 11 making airtight contact. Hold the victim's nose shut with the thumb and forefinger. Blow into the victim's mouth (gently in the case of infants) until his chest rises. Remove your mouth and release the hold on the victim's nose.



If you dislike direct contact, place a porous cloth between your mouth and the victim's.

If air cannot be blown in, check the position of the victim's head and jaw. Check the mouth for obstructions. Then try again blowing air more forcefully. If the chest still does not rise, turn the victim's face down and strike his back sharply to dislodge obstructions.

- 7 Let the victim exhale. Hear the out rush of air from the victim's mouth and nose.

Sometimes air enters the victim's stomach as evidenced by a swelling stomach. Expel the air by gently pressing the stomach during the exhalation period.

- 8 Repeat steps 6 and 7, eight to ten times rapidly. Then slow down to 10-12 times a minute. (20 times for infant)

Sometimes it may take hours for the victim to breathe normally. Continue giving artificial respiration till he recovers.

Mouth-to-nose method of artificial respiration

If the victim's mouth will not open, or has a blockage which you cannot clear then the mouth-to-nose method of providing artificial respiration should be followed. In this method use fingers of one hand to keep the victim's lips firmly shut. As shown in Fig 12 seal your lips around the victim's nostrils and blow air gently into his nose and suck back. Check if the victim's chest is rising and falling.

Fig 12



IT11021C

Artificial respiration in case of cardiac arrest

In the case of cardiac arrest where the heart has stopped beating, the following procedure has to be followed. One must act immediately as any delay will reduce the chances of the patient's recovery.

- 1 As shown in Fig 13, check if the carotid pulse in the neck can be felt. If the pulse is found feeble, go ahead with the following steps.

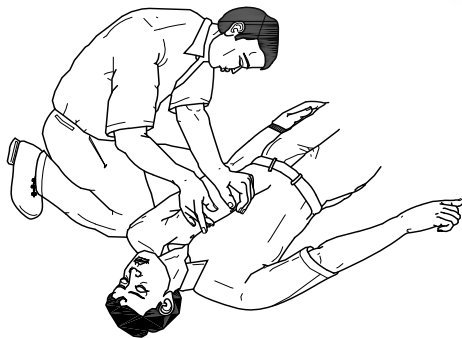
Fig 13



IT11021D

- 2 Lay the victim on his back on a firm surface. Kneel alongside as shown in Fig 14. Facing the chest, locate the lower part of the breastbone.

Fig 14



IT11021E

- 3 Place the palm of one hand on the centre of the lower part of the breastbone, keeping your fingers off the ribs. As shown in Fig 15, cover the palm with your other hand and lock your fingers together as shown.

Fig 15



IT11021F

- 4 Keeping your arms straight as shown in Fig 16, press sharply down on the lower part of the breastbone and release the pressure.

Fig 16



IT11021G

- 5 Repeat step 4, fifteen times at a rate of approximately one per second.
- 6 Recheck the carotid pulse as done in step 1, Fig 13.
- 7 Give two cycles of mouth-to-mouth artificial respiration as shown in Fig 17.

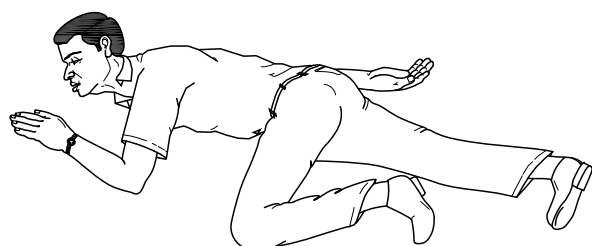
Fig 17



IT11021H

- 8 Give 10-15 more compressions of heart as done in step 4 and 6 follow it up with a further two cycles of mouth-to-mouth artificial respiration.
- 9 Checking for the carotid pulse. If the pulse is still feeble, repeat step 10 till the pulse improves.
- 10 As soon as the heartbeat returns, stop the compressions immediately and continue with mouth-to-mouth artificial respiration until natural breathing is fully restored.
- 11 Place the victim in the recovery position as shown in Fig.18. Keep him warm and get medical help quickly.

Fig 18



IT11021I

IMPORTANT POINTS TO NOTE AFTER GIVING ARTIFICIAL RESPIRATION TO VICTIMS

- 1 Even if the victim's breathing and heartbeat have recovered, do not delay in calling a doctor for a check up and treatment.
- 2 After the victim has recovered, keep the victim warm with a blanket, wrapped up with hot water bags. Stimulate circulation by stroking the insides of the arms and legs towards the heart.
- 3 When the victim revives, make him lie down. Do not let him exert himself as this may lead to a deterioration in his condition.
- 4 Do not give the victim any stimulant such as coffee, tea etc. until he is fully conscious.

ELECTRICAL SAFETY

Electric Shock

If a person happens to come in contact with an electrical live wire and if he has not insulated himself, then electric current flows through his body. Since the human body cannot withstand current flow more than a few tens of milliamps, the human body suffers a phenomenon generally known as **electric shock**. Electric shock may turn out to be hazardous to some of the parts of the human body and some times even to the life of the person.

The severity of an electric shock depends on:

- the level of current passing through the body
- how long does the current keep passing through the body.

Therefore, the higher the current or longer the time, the shock may result in a causality.

In addition to the above factors, other factors which influences the severity of shock are:

- age of the person receiving a shock
- surrounding weather condition
- condition of the floor (wet or dry)
- voltage level of electricity
- insulating property of the footwear or wet footwear, and so on.

Effects of electric shock

The effect of electric shock at very low voltage levels (less than 40 V) may only be an unpleasant tingling sensation. But this shock itself may be sufficient to cause someone to lose his balance and fall, resulting in casualty.

At higher voltage levels the muscles may contract and the person will be unable to break off from the contact by himself. He may lose consciousness. The muscles of the heart may contract spasmodically (fibrillation). This may even turn out to be fatal.

At an excessive level of voltage, the person receiving a shock may be thrown off his feet and will experience severe pain and possibly burns at the point of contact. This in most cases is fatal.

Electric shock can also cause burning of the skin at the point of contact.

Action to be taken in case of an electric shock

If the victim of an electric shock is in contact with the supply, break the contact the victim is making with the electricity by any one or more of the following.

- Switch off the electric power, insulate yourself and pull away the person from the electrical contact

OR

Remove the mains electric plug. Avoid direct contact with the victim. Wrap your hands using dry cloth or paper, if rubber gloves are not available.

OR

Remove the electric contact made by wrenching the cable/equipment/point free from contact using whatever is at hand to insulate yourself such as a wooden bar, rope, a scarf, the victim's coat-tails, any dry article of clothing, a belt, rolled up newspaper, non-metallic hose, PVC tubing, baked paper, tube etc. and break the contact by pushing or pulling the person or the cable/equipment/point free

OR

Stand on some insulating material such as dry wood, rubber or plastic, or whatever is at hand to insulate yourself and break the contact by pushing or pulling the person or the cable/equipment/point free.

If you are uninsulated, do not touch the victim with your bare hands. Otherwise you also will get a shock and become a victim.

If the victim is aloft (working on a pole or at raised place), take suitable measures to prevent him from falling or atleast ensure that his fall is safe.

Treatment to be given for the victim of electric shock

Electric burns on the victim may not look big/large. But it may be deep rooted. Cover the burnt area with a clean, sterile dressing. Get a doctor's help to treat him as quickly as possible.

If the victim is unconscious after an electric shock, but is breathing, carry out the following first aid:

- loosen the clothing at the neck, chest and waist
- place the victim in the recovery position as shown in Fig 18.
- Keep a constant check on the breathing and pulse rate. If you find them feeble, immediately give artificial respiration and press the lower rib to improve the heartbeat.
- Keep the casualty warm and comfortable.
- Send for a doctor immediately.

Do not give an unconscious person anything through the mouth.

Do not leave a unconscious person unattended.

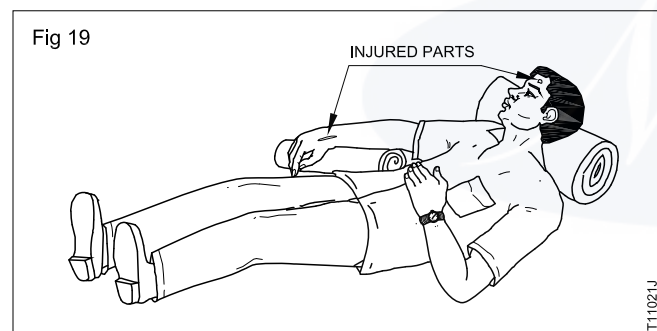
A person having received electric shock may also have burn injuries. DO NOT waste time by applying first aid to the burns until breathing has been restored and the patient can breathe normally unaided.

Treatment to be given in case of burns, severe bleeding

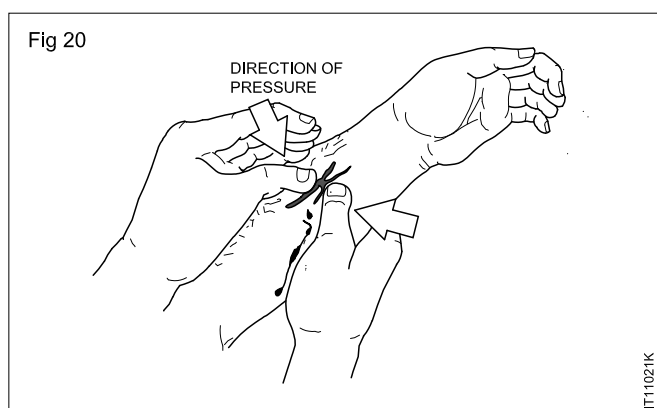
Burns caused due to electrical shock are very painful. If a large area of the body is burnt, clean the wound using clear water, or with clean paper, or a clean shirt. This treatment relieves the victim of pain. Do not give any other treatment on your own. Send for a doctor for further treatment.

A wound which is bleeding profusely, especially in the wrist, hand or fingers must be considered serious and must receive a doctor's attention. As an immediate first aid measure, carry out the following;

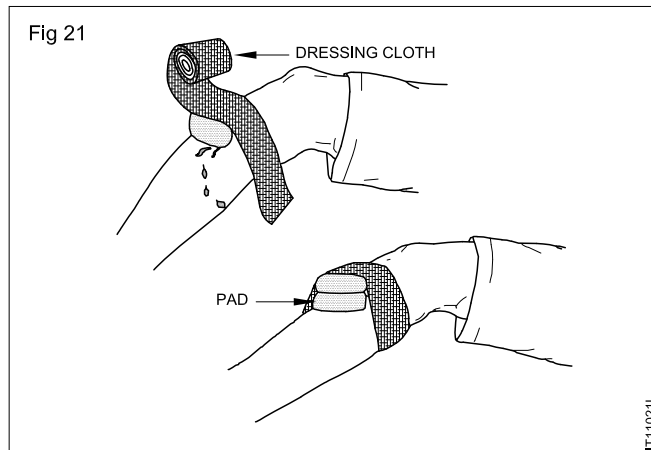
- make the patient lie down and rest
- if possible, raise the injured part above the level of the body as shown in Fig 19.



Squeeze together the sides of the wound as shown in Fig 20. Apply pressure as long as it is necessary to stop the bleeding.



When the bleeding stops temporarily, put a dressing over the wound using sterilized cotton, and cover it with a pad of soft material as shown in Fig 21.



If the wound is in the abdominal area (stab wound), caused by falling on a sharp tool, keep the patient bending over the wound to stop internal bleeding.

General procedural steps to be adopted for treating a person suffering from an electrical shock

- 1 Observe the situation. Choose the appropriate method(listed in earlier paragraphs) to release the person from electrical contact.

Do not run to switch off the supply that is far away or start searching for the mains switch.

- 2 Move the victim gently to the nearest ventilated place.
- 3 Check the victim's breathing and consciousness. Check if there are injuries in the chest or abdomen. Give artificial respiration/applying pressure on the heart if found necessary (refer in this lesson/exercise).

Use the most suitable method of giving artificial respiration depending upon the injuries if any on the chest/abdomen.

- 4 Send for a doctor.

Till the doctor arrives, you stay with the victim and render help as best as you can.

- 5 Place the victim in the recovery position.
- 6 Cover the victim with a coat, socks or any such thing to keep the victim warm.

Actions listed above must be taken systematically and briskly. Delay in treating the patient may endanger his life.

First - aid

Objectives: At the end of this lesson you shall be able to

- **state what is first aid**
 - **list the key aims of first aid**
 - **explain the ABC of the first aid**
 - **brief how to give first-aid for a victim who need first aid.**
-

First aid is defined as the immediate care and support given to an acutely injured or ill person, primarily to save life, prevent further deterioration or injury, plan to shift the victims to safer places, provide best possible comfort and finally help them to reach the medical centre/ hospital through all available means. It is an immediate life-saving procedure using all resources available within reach.

Imparting knowledge and skill through institutional teaching at younger age group in schools, colleges, entry point at industry level is now given much importance. Inculcating such habits at early age, helps to build good healthcare habits among people.

First aid procedure often consists of simple and basic life saving techniques that an individual performs with proper training and knowledge.

The key aims of first aid can be summarized in three key points:

- **Preserve life:** If the patient was breathing, a first aider would normally then place them in the recovery position, with the patient leant over on their side, which also has the effect of clearing the tongue from the pharynx. It also avoids a common cause of death in unconscious patients, which is choking on regurgitated stomach contents. The airway can also become blocked through a foreign object becoming lodged in the pharynx or larynx, commonly called choking. The first aider will be taught to deal with this through a combination of 'back slaps' and 'abdominal thrusts'. Once the airway has been opened, the first aider would assess to see if the patient is breathing.
- **Prevent further harm:** Also sometimes called prevent the condition from worsening, or danger of further injury, this covers both external factors, such as moving a patient away from any cause of harm, and applying first aid techniques to prevent worsening of the condition, such as applying pressure to stop a bleed becoming dangerous.
- **Promote recovery:** First aid also involves trying to start the recovery process from the illness or injury, and in some cases might involve completing a treatment, such as in the case of applying a plaster to a small wound.

Training

Basic principles, such as knowing to use an adhesive bandage or applying direct pressure on a bleed, are often acquired passively through life experiences. However,

to provide effective, life-saving first aid interventions requires instruction and practical training. This is especially true where it relates to potentially fatal illnesses and injuries, such as those that require cardiopulmonary resuscitation (CPR); these procedures may be invasive, and carry a risk of further injury to the patient and the provider. As with any training, it is more useful if it occurs before an actual emergency, and in many countries, emergency ambulance dispatchers may give basic first aid instructions over the phone while the ambulance is on the way. Training is generally provided by attending a course, typically leading to certification. Due to regular changes in procedures and protocols, based on updated clinical knowledge, and to maintain skill, attendance at regular refresher courses or re-certification is often necessary. First aid training is often available through community organization such as the Red cross and St. John ambulance.

ABC of first aid

ABC stands for airway, breathing and circulation.

- **Airway:** Attention must first be brought to the airway to ensure it is clear. Obstruction (choking) is a life-threatening emergency.
- **Breathing:** Breathing if stops, the victim may die soon. Hence means of providing support for breathing is an important next steps. There are several methods practiced in first aid.
- **Circulation:** Blood circulation is vital to keep person alive. The first aiders now trained to go straight to chest compressions through CPR methods.

When providing first aid one needs to follow some rule. There are certain basic norms in teaching and training students in the approach and administration of first aid to sick and injured.

Not to get panic

Panic is one emotion that can make the situation more worse. People often make mistake because they get panic. Panic clouds thinking and causes mistakes. First aider need calm and collective approach. If the first aider himself is in a state of fear and panic gross mistakes may result. It's far easier to help the suffering, when they know what they are doing, even if unprepared to encounter a situation. Emotional approach and response always lead to wrong doing and may cloud one to do wrong procedures. Hence be calm and focus on the given institution. Quick and confident approach can lessen the effect of injury.

Call medical emergencies

If the situation demands, quickly call for medical assistance. Prompt approach may save the life.

Surroundings play vital role

Different surroundings require different approach. Hence first aider should study the surrounding carefully. In other words, one need to make sure that they are safe and are not in any danger as it would be of no help that the first aider himself get injured.

Do no harm

Most often over enthusiastically practiced first aid viz. administering water when the victim is unconscious, wiping clotted blood (which acts as plug to reduce bleeding), correcting fractures, mishandling injured parts etc., would leads to more complication. Patients often die due to wrong FIRST AID methods, who may otherwise easily survive. Do not move the injured person unless the situation demands. It is best to make him lie wherever he is because if the patient has back, head or neck injury, moving him would causes more harm.

This does not mean do nothing. It means to make sure that to do something the care givers feel confident through training would make matters safe. If the first aider is not confident of correct handling it is better not to intervene or do it. Hence moving a trauma victim, especially an unconscious one, need very careful assessment. Removals of an embedded objects (Like a knife, nail) from the wound may precipitate more harm (e.g. increased bleeding). Always it is better to call for help.

Reassurance

Reassure the victim by speaking encouragingly with him.

Stop the bleeding

If the victim is bleeding, try to stop the bleeding by applying pressure over the injured part.

Golden hours

India have best of technology made available in hospitals to treat devastating medical problem viz. head injury, multiple trauma, heart attack, strokes etc, but patients often do poorly because they don't gain access to that technology in time. The risk of dying from these conditions, is greatest in the first 30 minutes, often instantly. This period is referred to as Golden period. By the time the patient reach hospitals, they would have passed that critical period. First aid care come handy to save lives. It helps to get to the nearest emergency room as quickly as possible through safe handling and transportation. The shorter that time, the more likely the best treatment applied.

Maintain the hygiene

Most importantly, first aider need to wash hands and dry before giving and first aid treatment to the patient or wear gloves in order to prevent infection.

Cleaning and dressing

Always clean the wound thoroughly before applying the bandage lightly wash the wound with clean water.

Not to use local medications on cuts or open wounds

They are more irritating to tissue than it is helpful. Simple dry cleaning or with water and some kind of bandage are best.

CPR (Cardio-Pulmonary Resuscitation) can be life-sustaining

CPR can be life sustaining. If one is trained in PR and the person is suffering from choking or finds difficulty in breathing, immediately begin CPR. However, if one is not trained in CPR, do not attempt as you can cause further injury. But some people do it wrong. This is a difficult procedure to do in a crowded area. Also there are many studies to suggest that no survival advantage when bystanders deliver breaths to victims compared to when they only do chest compressions. Second, it is very difficult to carry right maneuver in wrong places. But CPR, if carefully done by highly skilled first aiders is a bridge that keeps vital organs oxygenated until medical team arrives.

Declaring death

It is not correct to declare the victim's death at the accident site. It has to be done by qualified medical doctors.

How to report an emergency?

Reporting an emergency is one of those things that seems simple enough, until actually when put to use in emergency situations. A sense of shock prevail at the accident sites. Large crowd gather around only with inquisitive nature, but not to extend helping hands to the victims. This is common in road side injuries. No passer-by would like to get involved to assist the victims. Hence first aid management is often very difficult to attend to the injured persons. The first aiders need to adapt multi-task strategy to control the crowd around, communicate to the rescue team, call ambulance etc., all to be done simultaneously. The mobile phones helps to a greater deal for such emergencies. Few guidelines are given below to approach the problems.

Assess the urgency of the situation. Before you report an emergency, make sure the situation is genuinely urgent. Call for emergency services if you believe that a situation is life-threatening or otherwise extremely disruptive.

- A crime, especially one that is currently in progress. If you're reporting a crime, give a physical description of the person committing the crime.
- A fire - If you're reporting a fire, describe how the fire started and where exactly it is located. If someone has already been injured or is missing, report that as well.
- A life-threatening medical emergency, explain how the incident occurred and what symptoms the person currently displays.
- A car crash - Location, serious nature of injuries, vehicle's details and registration, number of people involved etc.

Call emergency service

The emergency number varies - 100 for Police & Fire, 108 for Ambulance.

Report your location

The first thing the emergency dispatcher will ask is where you are located, so the emergency services can get there as quickly as possible. Give the exact street address, if you're not sure of the exact address, give approximate information.

Give the dispatcher your phone number

This information is also imperative for the dispatcher to have, so he or she is able to call back if necessary.

Describe the nature of the emergency

Speak in a calm, clear voice and tell the dispatcher why you are calling. Give the most important details first, then answer the dispatcher's follow-up question as best you can.

Do not hang up the phone until you are instructed to do so. Then follow the instructions you were given.

How to do basic first aid?

Basic first aid refers to the initial process of assessing and addressing the needs of someone who has been injured or is in physiological distress due to choking, a heart attack, allergic reactions, drugs or other medical emergencies. Basic first aid allows one to quickly determine a person's physical condition and the correct course of treatment.

Important guideline for first aiders

Evaluate the situation

Are there things that might put the first aider at risk. When faced with accidents like fire, toxic smoke, gasses, an unstable building, live electrical wires or other dangerous scenario, the first aider should be very careful not to rush into a situation, which may prove to be fatal.

Remember A-B-Cs

The ABCs of first aid refer to the three critical things the first aiders need to look for.

- Airway - Does the person have an unobstructed airway?
- Breathing - Is the person breathing?
- Circulation - Does the person show a pulse at major pulse points (wrist, carotid artery, groin)

Avoid moving the victim

Avoid moving the victim unless they are in immediate danger. Moving a victim will often make injuries worse, especially in the case of spinal cord injuries.

Call emergency services

Call for help or tell someone else to call for help as soon as possible. If alone in at the accident scene, try to establish breathing before calling for help, and do not leave the victim alone unattended.

Determine responsiveness

If a person is unconscious, try to rouse them by gently shaking and speaking to them.

If the person remains unresponsive, carefully roll them on the side (recovery position) and open his airway.

- Keep head and neck aligned.
- Carefully roll them onto their back while holding his head.
- Open the airway by lifting the chin.

Fig 1



FT110311

Look, listen and feel for signs of breathing

Look for the victim's chest to raise and fall, listen for sounds of breathing.

If the victim is not breathing, see the section below

- If the victim is breathing, but unconscious, roll them onto their side, keeping the head and neck aligned with the body. This will help drain the mouth and prevent the tongue or vomit from blocking the airway.

Check the victim's circulation

Look at the victim's colour and check their pulse (the carotid artery is a good option; it is located on either side of the neck, below the jaw bone). If the victim does not have a pulse, start CPR.

Treat bleeding, shock and other problems as needed

After establishing that the victim is breathing and has a pulse, next priority should be to control any bleeding. Particularly in the case of trauma, preventing shock is the priority.

- **Stop bleeding:** Control of bleeding is one of the most important things to save a trauma victim. Use direct pressure on a wound before trying any other method of managing bleeding.
- **Treat shock:** Shock, a loss of blood flow to the body, frequently follows physical and occasionally psychological trauma. A person in shock will frequently have ice cold skin, be agitated or have an altered mental status, and have pale colour to the skin around the face and lips. Untreated, shock can be fatal. Anyone who has suffered a severe injury or life-threatening situation is at risk for shock.
- **Choking victim:** Choking can cause death or permanent brain damage within minutes.
- **Treat a burn:** Treat first and second degree burns by immersing or flushing with cool water. Don't use creams, butter or other ointments, and do not pop blisters. Third degree burns should be covered with a damp cloth. Remove clothing and jewellery from the burn, but do not try to remove charred clothing that is stuck to burns.
- **Treat a concussion:** If the victim has suffered a blow to the head, look for signs of concussion. Common symptoms are: loss of consciousness following the injury, disorientation or memory impairment, vertigo, nausea, and lethargy.
- **Treat a spinal injury victim:** If a spinal injury is suspected, it is especially critical, not move the victim's head, neck or back unless they are in immediate danger.

Stay with the victim until help arrives

Try to be a calming presence for the victim until assistance can arrive.

Unconsciousness (COMA)

Unconscious also referred as Coma, is a serious life threatening condition, when a person lie totally senseless and do not respond to calls, external stimulus. But the basic heart, breathing, blood circulation may be still intact, or they may also be failing. If unattended it may lead to death.

The condition arises due to interruption of normal brain activity. The causes are too many.

- Shock (Cardiogenic, Neurogenic)
- Head injury (Concussion, Compression)
- Asphyxia (obstruction to air passage)
- Extreme of body temperature (Heat, Cold)
- Cardiac arrest (Heart attack)
- Stroke (Cerebro-vascular accident)
- Blood loss (Haemorrhage)
- Dehydration (Diarrhoea & vomiting)
- Diabetes (Low or high sugar)
- Blood pressure (Very low or very high)
- Over dose of alcohol, drugs
- Poisoning (Gas, Pesticides, Bites)
- Epileptic fits (Fits)
- Hysteria (Emotional, Psychological)

The following symptoms may occur after a person has been unconscious:

- Confusion
- Drowsiness
- Headache
- Inability to speak or move parts of his or her body (see stroke symptoms)
- Light headedness
- Loss of bowel or bladder control (incontinence)
- Rapid heartbeat (palpitation)
- Stupor

First aid

- Call EMERGENCY number.
- Check the person's airway, breathing, and pulse frequently. If necessary, begin rescue breathing and CPR.
- If the person is breathing and lying on the back and after ruling out spinal injury, carefully roll the person onto the side, preferably left side. Bend the top leg so both hip and knee are at right angles. Gently tilt the head back to keep the airway open. If breathing or pulse stops at any time, roll the person on to his back and begin CPR.
- If there is a spinal injury, the victim's position may have to be carefully assessed. If the person vomits, roll the entire body at one time to the side. Support the neck and back to keep the head and body in the same position while you roll.
- Keep the person warm until medical help arrives.
- If you see a person fainting, try to prevent a fall. Lay the person flat on the floor and raise the level of feet above and support.

- If fainting is likely due to low blood sugar, give the person something sweet to eat or drink when they **BECOME CONSCIOUS.**

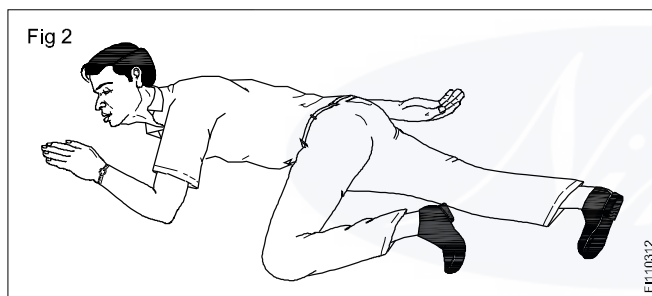
DO NOT

- Do not give an unconscious person any food or drink.
- Do not leave the person alone.
- Do not place a pillow under the head of an unconscious person.
- Do not slap an unconscious person's face or splash water on the face to try to revive him.

Loss of consciousness may threaten life if the person is on his back and the tongue has dropped to the back of the throat, blocking the airway. Make certain that the person is breathing before looking for the cause of unconsciousness. If the injuries permit, place the casualty in the recovery position with the neck extended. Never give anything by mouth to an unconscious casualty.

How to diagnose an unconscious injured person

- Consider alcohol: look for signs of drinking, like empty bottles or the smell of alcohol.



- Consider epilepsy: are there signs of a violent seizure, such as saliva around the mouth or a generally dishevelled scene?
- Think insulin: might the person be suffering from insulin shock (see 'How to diagnose and treat insulin shock')?
- Think about drugs: was there an overdose? Or might the person have under dosed - that is not taken enough of a prescribed medication?
- Consider trauma: is the person physically injured?
- Look for signs of infection: redness and/ or red streaks around a wound.
- Look around for signs of Poison: an empty bottle of pills or a snakebite wound.
- Consider the possibility of psychological trauma: might the person have a psychological disorder of some sort?
- Consider stroke, particularly for elderly people.
- Treat according to what you diagnose.

Shock

A severe loss of body fluid will lead to a drop in blood pressure. Eventually the blood's circulation will deteriorate and the remaining blood flow will be directed to the vital organs such as the brain. Blood will therefore be directed away from the outer area of the body, so the victim will appear pale and the skin will feel ice cold.

Fig 3



As blood flow slows, so does the amount of oxygen reaching the brain. The victim may appear to be confused, weak, and dizzy and may eventually deteriorate into unconsciousness. Try to compensate for this lack of oxygen, the heart and breathing rates both speed up, gradually becoming weaker, and may eventually cease.

Potential causes of shock include: severe internal or external bleeding; burns; severe vomiting and diarrhea, especially in children and the elderly; problems with the heart.

Symptoms of shock

Victims appear pale, ice cold, pulse appear initially faster and gets slower, breathing become shallow. Weakness, dizziness, confusion continue. If unattended the patient may become unconscious and die.

Shock kills, so it is vital that you can recognize these signs and symptoms. With internal bleeding in particular, shock can occur sometime after an accident, so if a person with a history of injury starts to display these symptoms coupled with any of the symptoms of internal bleeding, advise her to seek urgent medical attention. Or take or send her to hospital.

First aid

Keep the patient warm and at mental rest. Assure of good air circulation and comfort. Call for help to shift the patient to safer place/ hospital.

- **Warmth:** Keep the victim warm but do not allow her to get overheated. If you are outside, try to get something underneath her if you can do easily. Wrap blankets and coats around her, paying particular attention to the head, through which much body heat is lost.
- **Air:** Maintain careful eye on the victim's airway and be prepared to turn her into the recovery position if

necessary, or even to resuscitate if breathing stops. Try to keep back bystanders and loosen tight clothing to allow maximum air to victim.

- **Rest:** Keep the victim still and preferably sitting or lying down. If the victim is very giddy, lay her down with her legs raised to ensure that maximum blood and therefore maximum oxygen is sent to the brain.

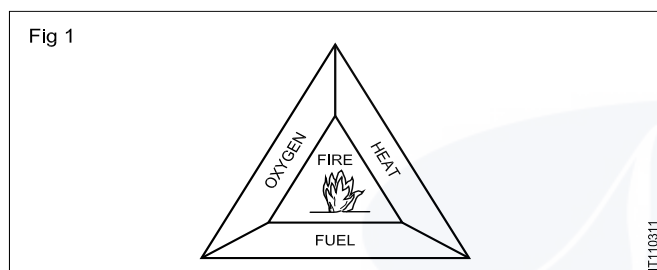
Practice Extinguishing fire

Objectives: At the end of this lesson you shall be able to

- state the effects of a fire break out
- state the conditions required for combustion relevant to fire prevention
- state the general precautionary measures to be taken for fire prevention
- determine the correct type of fire extinguisher required for a particular function.

Fire

Fire is nothing but burning of a combustible material. For combustion the three main requirements are shown in Fig 1.



Fuel

Fuel can be any combustible substance in the form of a solid, liquid or gas. Examples; wood, paper, petrol, kerosene, LPG etc., The fuel will catch fire and burn provided a high enough temperature(heat) is brought about and a continuous supply of oxygen is given. It is important to note that without fuel, combustion cannot take place.

Heat

Fuels will begin to burn at a certain temperature. Different types of fuels need different temperatures to catch fire and burn. For example, wood needs a higher temperature to catch fire and burn than paper. Petrol needs much lesser temperature to catch fire and burn than paper. Generally liquid fuels give off vapour when heated. It is this vapour which ignites. Some liquids such as petrol do not have to be heated as they give off vapour at room temperature(15°C - 25°C) itself. It is important to note that without heat, fuel cannot get ignited(catch fire) and hence combustion cannot take place.

Oxygen

Oxygen exists in air. The amount of oxygen in air is sufficient to continue the combustion once it occurs. Hence to keep a fire burning, oxygen is a must. It is important to note that without oxygen, combustion cannot continue to take place.

Controlled and uncontrolled fire

Fire is a boon to mankind. Without fire, there would not be cooked food or hot water for bath as and when we want it. At the same time if the fire does not get constrained to a place of requirement, fire can become a bane(curse) to mankind. An uncontrolled fire can cause such a disaster which not only leads to destruction of material but also endanger the life of persons. Hence, the lesson one must never forget is, *keep the fire under control*. Every effort must be made to prevent uncontrolled fire. When there is a fire outbreak, it must be controlled and extinguished immediately without any delay.

Preventing fire

The majority of fires begin with small outbreaks. If this is not noticed, fire goes out of control and will be on its way of destruction. Hence, most fires could be prevented if suitable care is taken by following some simple common sense rules as given below.

- Do not accumulate combustible refuse such as cotton waste, waste or cloth soaked with oil, scrap wood, paper, etc. in odd corners. These refuse should be in their collection bins or points.
- Do not misuse or neglect electrical equipments or electrical wiring as this may cause electrical fire. Loose connections, low rated fuses, overloaded circuits causes over heating which may in turn lead to fire. Damaged insulation between conductors in cables cause electrical short circuit and cause fire.
- Keep away clothing and other materials which might catch fire from heating appliances. Make sure the soldering iron is disconnected from power supply and is kept safe in its stand at the end of the working day.
- Store highly flammable liquids and petroleum mixtures such as thinner, adhesive solutions, solvents, kerosene, spirit, LPG gas etc. in the storage area exclusively meant for storage of flammable materials.
- Turn off blowlamps and torches when they are not in use.

Controlling and Extinguishing fire

Isolating or removing any of three factors illustrated in Fig 1, will control and extinguish fire. There are three basic ways of achieving this.

1 Starving the fire of fuel

To remove the fuel which is burning or cut further supply of fuel to the fire.

2 Smothering

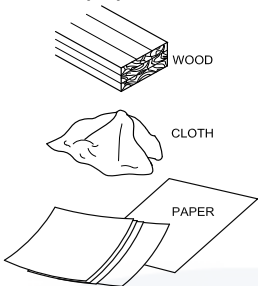
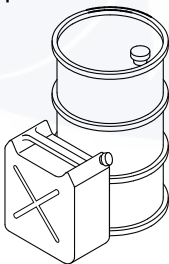
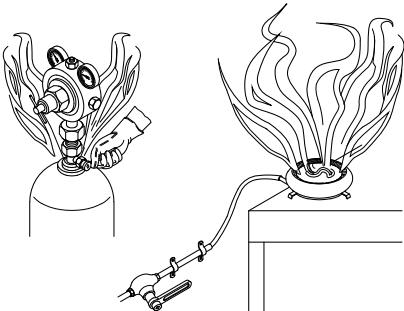
To stop the supply of oxygen to the fire by blanketing the fire with foam, sand etc.

3 Cooling

To reduce the temperature of the fire by spraying water and thus cooling the fire.

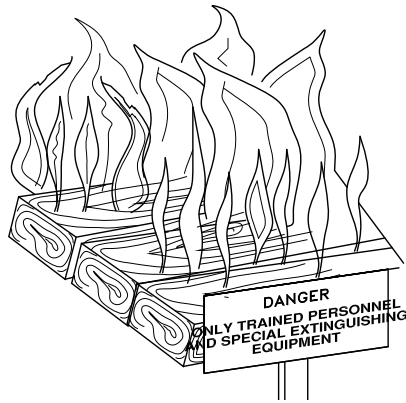
By any one of the above three methods, fire can be first controlled and then extinguished.

For the purpose of determining the best method of extinguishing different types of fires, fires are classified under four main classes based on the type of fuel as given in Table 1.

TABLE 1		
Classification of Fire	Fuel involved	Precautions and extinguishing
Class A Fire	Wood, paper cloth etc. 	Most effective method is cooling with water.
Class B Fire	Solid materials Flammable liquids & liquefiable solids 	Jets of water should be sprayed on the base Should be smothered. The aim is to cover the entire surface of the burning liquid. This has the effect of cutting off the supply of oxygen to the fire. Water should never be used on burning liquids. Foam, dry powder or CO ₂ may be used on this type of fire.
Class C Fire	Gas and liquified gas 	Extreme caution is necessary in dealing with liquefied gases. There is a risk of explosion and sudden outbreak of fire in the entire vicinity. If an appliance fed from a cylinder catches fire - shut off the supply of gas. The safest course is to raise an alarm and leave the fire to be dealt with by trained personnel. Dry powdered extinguishers are used on this type of fire.

Class D Fire

Involving metals



The standard range of fire extinguishing agents is inadequate or dangerous when dealing with metal fires.

Fire in electrical equipment:

Carbon dioxide, dry powder, and vapourising liquid (CTC) extinguishers can be used to deal with fires in electrical equipment.

Foam or liquid (eg. water) extinguishers must not be used on electrical equipment at all.

Fire extinguishers

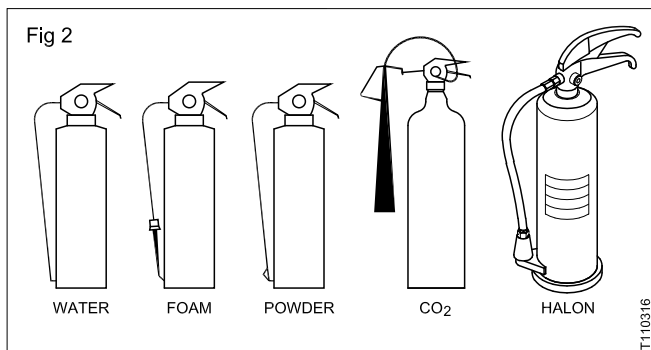
Different fire extinguishing agents should be used for different types of fires as listed in Table 1. Using a wrong type of extinguishing agent can make things worse.

A fire extinguishing agent is the material or substance used to put out the fire. These extinguishing materials are usually (but not always) contained in a container called the 'fire extinguisher' with a mechanism for spraying into the fire when needed.

There is no classification for **electrical fires** as such as these are only fires in materials where electricity is present. To control electrical fire in a building the electrical supply should be cut off first.

Types of fire extinguishers

Many types of fire extinguishers are available with different extinguishing *agents* to deal with different classes of fires as shown in Fig 2. Always check the operating instructions on the extinguisher before use.

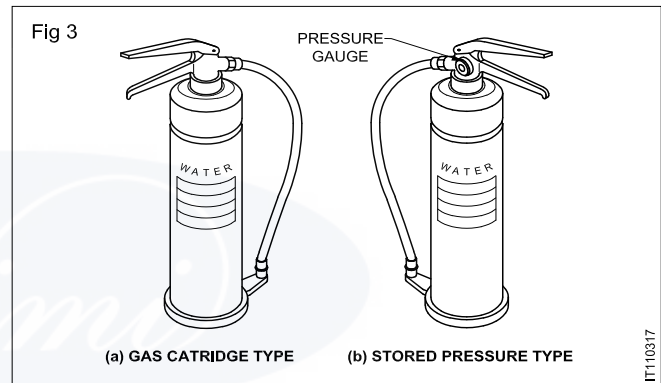


Water-filled extinguishers

In water-filled extinguishers, as shown in Fig 3, there are two types based on the method of operating the extinguisher.

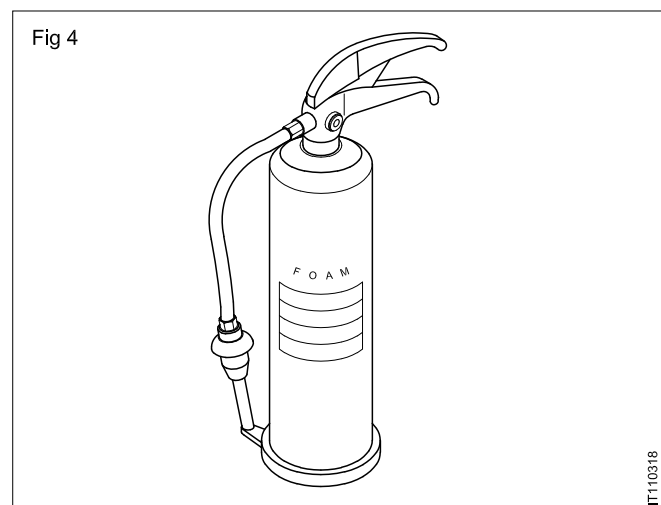
- a) Cartridge type
- b) Stored pressure type

In both the methods of operation, the discharge can be interrupted as required. This is to conserve the contact area and to prevent unnecessary damage to the material due to water.



Foam extinguishers

These may be stored pressure or gas cartridge types as shown in Fig 4.



Most suitable for:

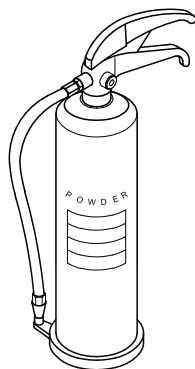
- flammable liquid fires
- running liquid fires.

Not to be used in fires where electrical equipment is involved.

Dry powder extinguishers

Extinguishers fitted with dry powder may be of the gas cartridge or stored pressure type as shown in Fig 5. Appearance and the method of operation is the same as that of water-filled one. The main distinguishing feature is the fork-shaped nozzle. Powders have been specially developed to deal with Class D fires.

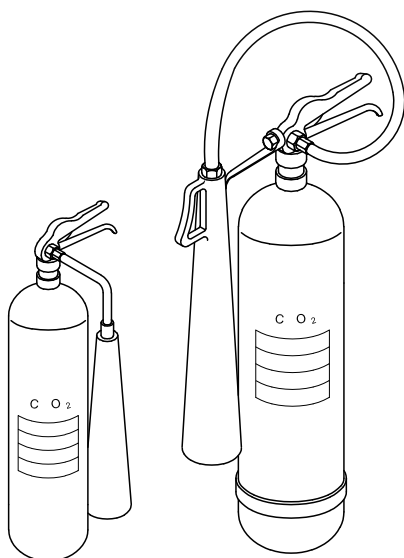
Fig 5



Carbondioxide (CO₂)

This type is easily distinguished by the distinctively shaped discharge horn as shown in Fig 6. These extinguishers are suitable for fires on flammable liquids and liquefiable solids. Best suited where contamination by deposits must be avoided. Not generally effective in the open air.

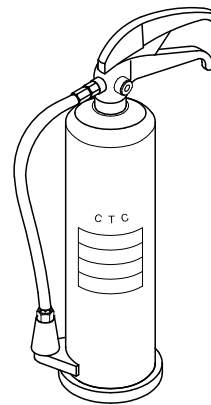
Fig 6



Halon Extinguishers (Fig 7)

Carbontetrachloride(CTC) and Bromochlorodifluoro methane (BCF). They may be either gas cartridge or non-conductive.

Fig 7



The fumes given off by these extinguishers are dangerous especially in confined space.

General procedure to be adopted in the event of a fire

- 1 Raise a loud alarm by using any of the following.

Adopt any one method of giving an alarm signal for fire breaking in your institute/ workshop.

- Raising your voice and shouting Fire! Fire!Fire! to call the attention of others.
- Running towards the fire shouting Fire! Fire! and actuate fire alarm/bell/siren. This alarm/bell/siren to be actuated only in case of fire.
- Any other means by which the attention of others can be called and are made to understand there is a fire break out.

- 2 On receipt of the fire alarm signal, do the following:

- stop the normal work you are doing
- turn off the power for all machinery and equipments
- switch off fans/air circulators/exhaust fans
- switch off the mains if accessible.

- 3 If you are not involved in fire fighting team, then,

- evacuate the working premises
- close the doors and windows, but do not lock or bolt
- assemble at a safe open place along with the others
- if you are in the room/place where the fire has broken out, leave the place calmly through the emergency exit.

- 4 If you are involved in the fire fighting team,

- take instructions/give instructions for an organized way of fighting the fire.

If you are taking instructions,

- follow the instructions systematically. Do not panic. Do not get trapped in fire or smoke in a hurry.

If you are giving instructions,

- assess the class of fire(class A,B,C or D)
 - send for sufficient assistance and fire brigade
 - judge the magnitude of the fire. Locate locally available suitable means to put-out the fire.
 - ensure emergency exit paths are clear of obstructions. Attempt to evacuate the people and explosive materials, substances that can serve as further fuel for fire within the vicinity of the fire break.
 - Allot clear activity to persons involved in firefighting by name to avoid confusion.
- Control and extinguish the fire using the right type of fire extinguisher and making use of the available assistance effectively.
- 5 After fully extinguishing the fire, make a report of the fire accident and the measures taken to put out the fire, to the authorities concerned.

Reporting all fires however small they are, helps in the investigation of the cause of the fire. It helps in preventing the same kind of accident occurring again.



Space for recording additional information



Basic electricity, fuses and switches

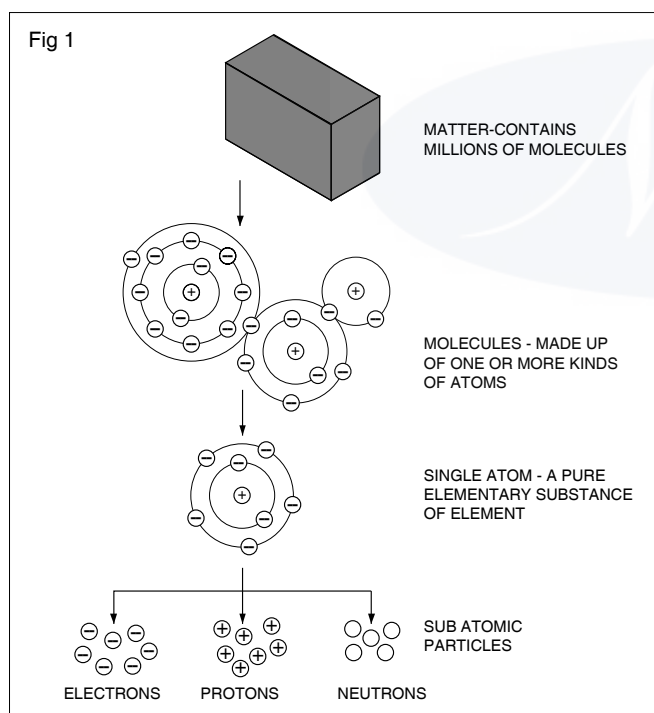
Objectives: At the end of this lesson you shall be

- draw the structure of an atom and name its sub-atomic particles
- identify conductors and insulators from their atomic structure of a given material
- list the commonly used conductors, insulators and semiconductors
- explain the meaning of voltage, its unit and methods
- explain the meaning of electrical current
- explain the meaning of electrical resistance and its unit.

Matter, atomic structure

Anything that one can see, feel, or use is matter. Matter is anything that has weight and occupies space. Matter can be in the form of a solid like stone, wood, metal or liquid like water, alcohol or gas like oxygen, hydrogen, carbon dioxide etc.

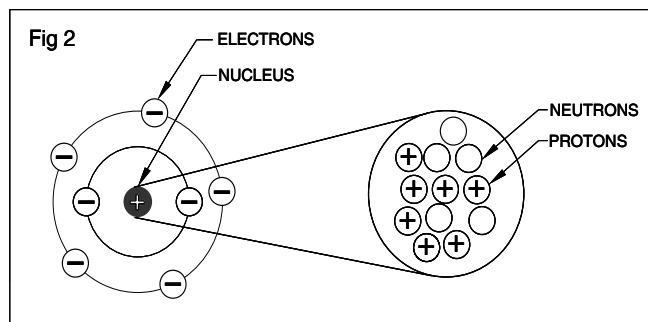
Matter is made up of tiny particles called molecules. Molecules are made up of even smaller particles called atoms. Atoms are made up of further smaller particles called **protons**, **neutrons** and **electrons** as shown in Fig 1.



Atomic structure

Although an atom cannot be seen through naked eyes, Neils Bohr has suggested the structure of an atom as shown in Fig 2 which is universally accepted.

As shown in Fig 2, **nucleus** is the central part of the atom. Nucleus is made up of **protons** which are positively



charged and **neutrons** which do not have any charge. Electrons revolve around the nucleus and are negatively charged. Generally the number of electrons(-ve charge) in an atom will be equal to the number of protons(+ve charge). The number of protons in the nucleus of an atom of one element differs from that of an atom in another element. Table below gives the number of Protons in atoms of some elements.

Name of the element	No. of protons	Atomic number
Hydrogen	1	1
Oxygen	8	8
Silver	47	47
Gold	79	79

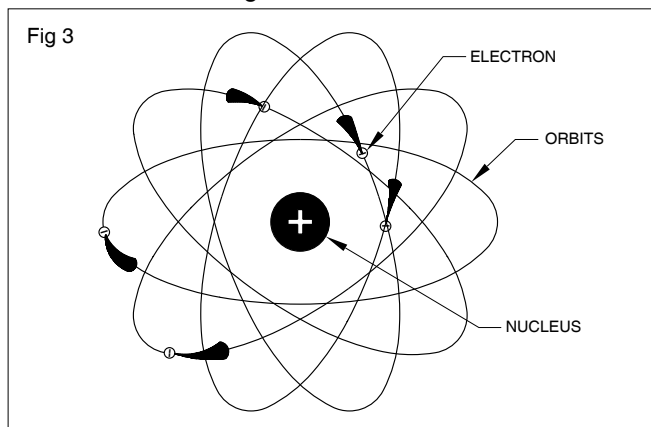
Because the number of electrons in an atom is equal to the number of protons, an atom will be electrically neutral when it is not participating in any chemical, electrical or other activity. In other words, the positive charges of protons will neutralise the negative charges of electrons resulting in zero net charge of an atom.

Any atom is recognised by the following specifications;

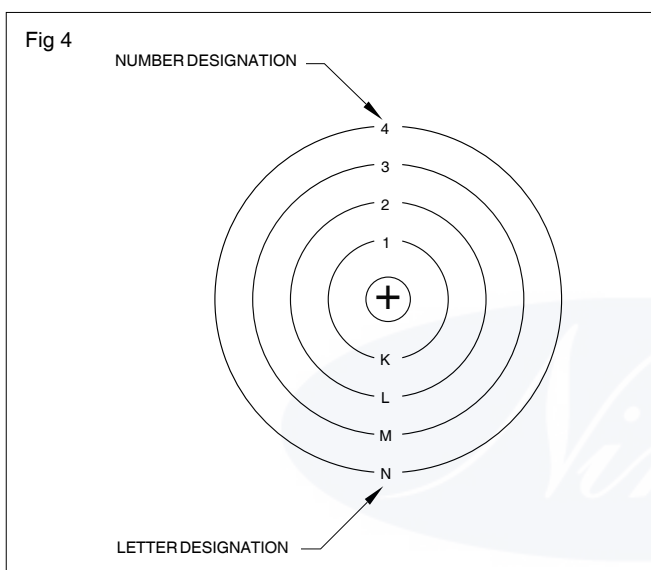
Atomic number - number of protons or number of electrons.

Atomic weight - total number of protons and neutrons.

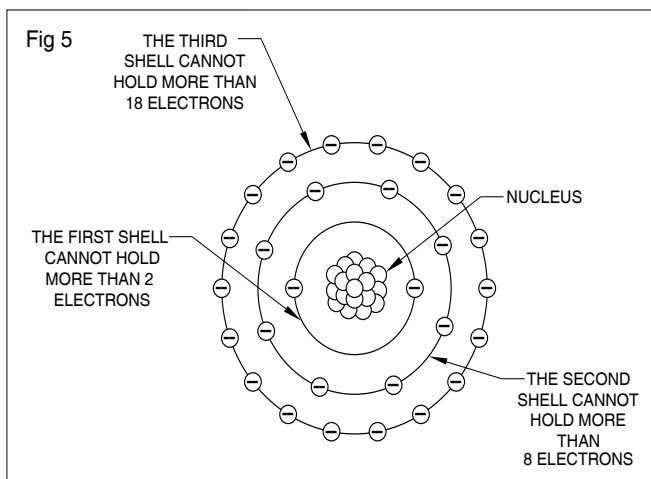
In all atoms, electrons revolve round the nucleus in different orbits as shown in Fig 3.



These orbits are also called shells. Shells are identified by numbers or by letters as shown in Fig 4.



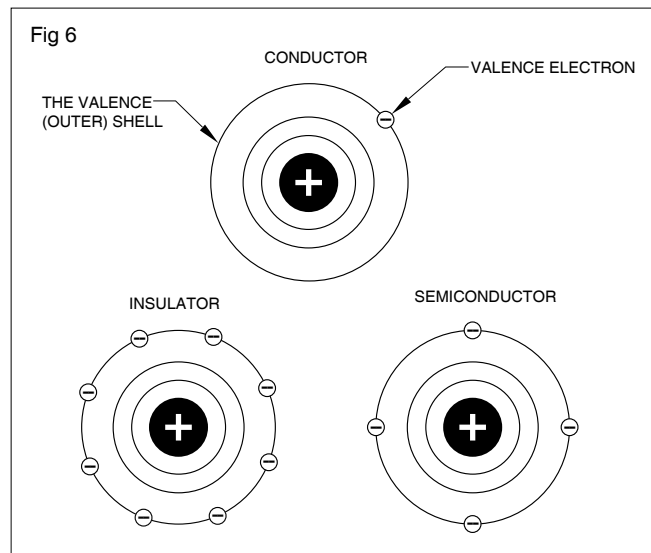
The maximum number of electrons in each orbit is given by $2n^2$ where n is the orbit number as shown in Fig 5.



As the protons in the nucleus are positively charged, the electrons revolving round the nucleus will have a force of attraction towards the nucleus (opposite charges attract each other). This force is maximum for the electrons in the innermost shell and minimum for the outermost shell because of its distance from the nucleus. (the magnitude of attraction is inversely proportional to the distance between the charges)

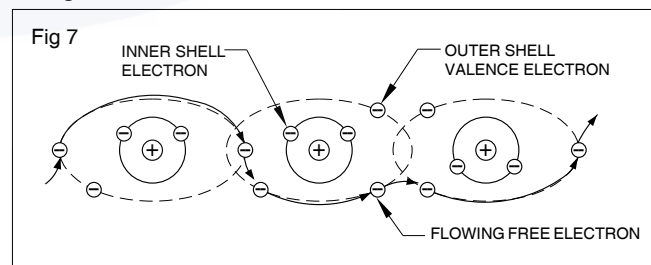
The outermost orbit of an atom is called the **valence orbit**. The electrons in this orbit are called **valence electrons**. These valence electrons are loosely bound to the nucleus as they are far from nucleus.

Depending upon the number of valence electrons in an atom, the materials can be classified as conductors, insulators and semiconductors as shown in Fig 6.



CONDUCTORS

In atoms having one or two electrons in its valence orbit, the electrons can be pulled out of their orbit and made available as free electrons. Such materials are called conductors. With the application of external force, these electrons become free to participate in conduction. Electricity is defined as the flow of electrons as shown in Fig 7.



Copper atom has one valence electron. Hence, copper is a very good conductor of electricity. In other words the conductivity of copper is very good. Given below is a list of metal conductors, in the order of their conductivity.

- 1 Silver
- 2 Copper
- 3 Gold
- 4 Aluminium
- 5 Iron

From the above list, it is evident that silver is the best conductor followed by copper, gold and aluminium. Copper and aluminium are available in abundance in nature, hence the cost of these metals is cheaper than the other conductors. Therefore copper and aluminium are the most commonly used for electrical work. Gold is used for special

purposes as conductor as this metal has certain special property compared to copper and silver.

INSULATORS

If the number of valence electrons in an atom is 5, 6, 7 or 8 then they tend to bind to the atom. Therefore even if a large external energy is applied, none of the electrons or very few electrons come out of their orbit and be available as free electrons. Hence such materials which are not good conductors of electricity are called non-conductors or insulators. Some common insulators are listed below, in the descending order of non-conductivity.

1 Porcelain	5 Paper
2 Glass	6 Air
3 Plastic	7 Wood
4 Rubber	

SEMICONDUCTORS

If the number of valence electrons in an atom is 4, materials

composed of such atoms are called semiconductors. These have the characteristics in-between that of a conductor and an insulator. The conductivity of semiconductors also depends upon the temperature in which they are used. At low temperatures semiconductors behaves as insulators. At high temperatures they behave as fairly good conductors. Examples of most used semiconducting materials are silicon and germanium.

CURRENT

The motion of electrons is called **electric conduction or electric current**. When no external energy is supplied, the free electrons in a conductor move from atom to atom within the conductor. But this motion is at random, without any organised flow through the conductor. Hence when the conductor is considered as a whole, from one end to other end, the current is zero. However when an external force is given to these random moving free electrons, then the motion of electrons will be organised, resulting in the movement of electrons in a particular direction through the wire. This is called flow of electric current.

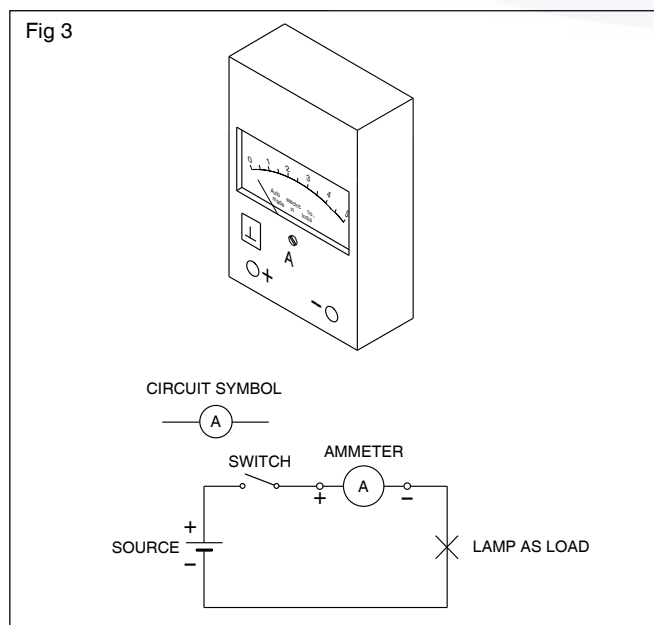
Simple electrical circuit and its elements

Objectives: At the end of this lesson you shall be able to

- explain the emf, potential difference, their units and method of measurement (Voltmeter)
- explain the current, its units and method of measurement (Ammeter)
- explain resistance and its unit.

Ampere

The unit of current (abbreviated as I) is ampere (symbol A). If 6.24×10^{18} electrons pass through a conductor per second, then we can say one ampere current has passed through the conductor.



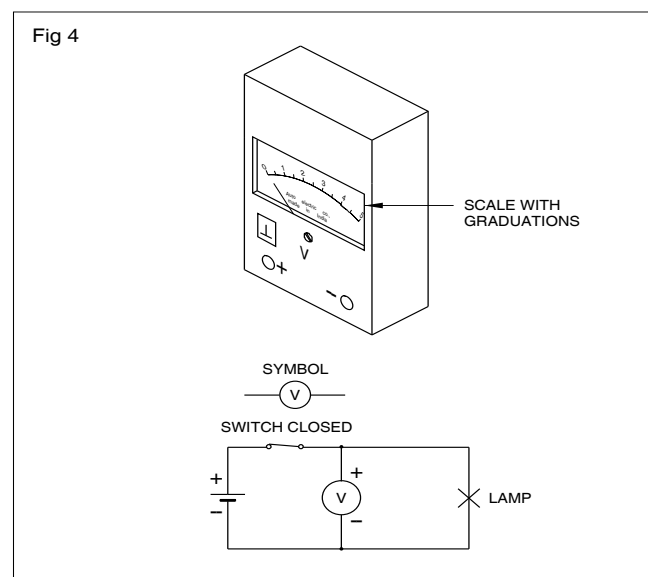
The terminals of the battery are indicated in the circuit symbol by two lines, the longer line for the **positive** and the shorter for the **negative** terminal.

Within the battery the negative terminal contains an excess of electrons whereas the positive terminal has a

deficit of electrons. The battery is said to have an electromotive force (emf) which is available to drive the free electrons in the closed path of the electrical circuit. The difference in the distribution of electrons between the two terminals of the battery produces this emf.

Volt

The unit of electromotive force is the volt (symbol V) and the emf is commonly referred as 'voltage'. When the battery is connected to any load, the voltage measured across the terminals is called potential difference (PD) and this will be slightly less than the value of emf.



Resistance

In addition to the current and voltage there is a third quantity which plays a role in a circuit, called the electrical resistance. Resistance is the property of a material by which it opposes the flow of electric current.

Ohm

The unit of electrical resistance (abbreviated as R) is ohm (symbol Ω).

For the decimal multiples or decimal sub-multiples of the ohm we use the following expressions:

$$1 \text{ megohm} = 1 \text{ M } \Omega = 1000000 \Omega = 1 \times 10^6 \Omega$$

$$1 \text{ kilo-ohm} = 1 \text{ k } \Omega = 1000 \Omega = 1 \times 10^3 \Omega$$

$$1 \text{ milli-ohm} = 1 \text{ m } \Omega = 1/1000 \Omega = 1 \times 10^{-3} \Omega$$

$$1 \text{ micro-ohm} = 1 \text{ } \mu \Omega = 1/1000000 \Omega = 1 \times 10^{-6} \Omega$$

Space for recording additional information



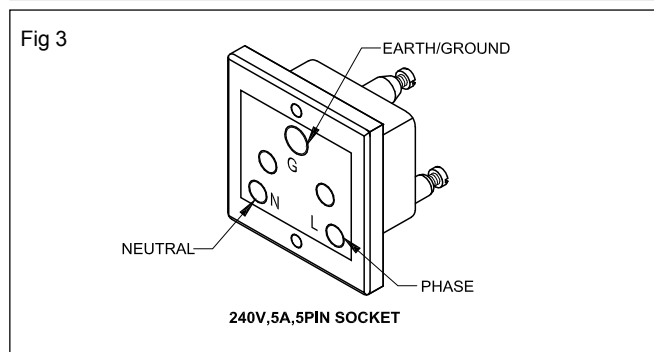
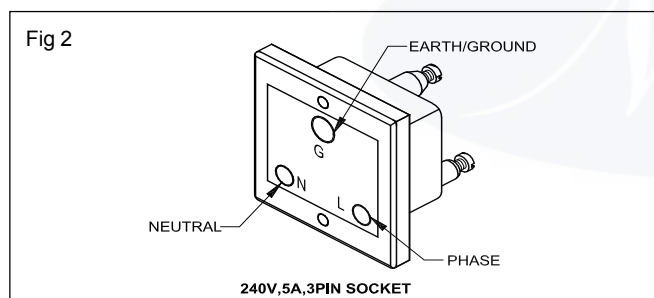
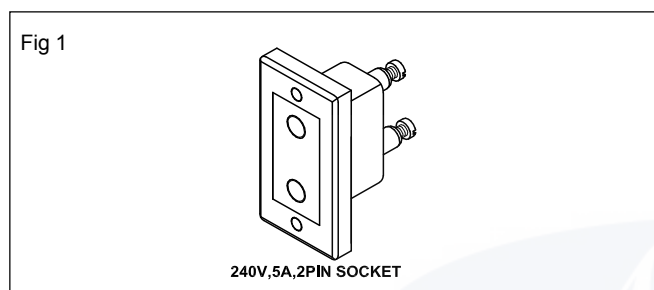
Concept of Fuses & switches

Objectives : At the end of this lesson you shall be able to

- state the functions, types and specifications of electrical accessories
- state the necessity of a fuse.

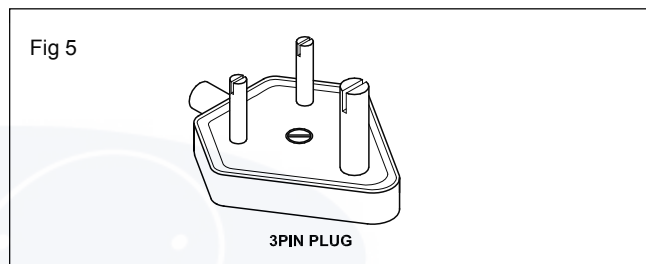
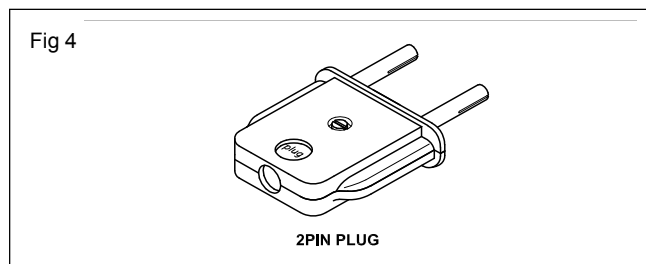
Sockets

Function of the socket is to keep the AC main supply made available and it can be used whenever required. The types of the sockets are normally 2 pins, 3 pins and 5 pins sockets. In that 2 pin socket will not be having earth point but 3 pin and 5 pin sockets will have earth point. The purpose of providing earth point is to give easiest path for leakage current on the metal body of electrical apparatus/equipment such as heaters, iron boxes, radios, tape recorders and computers etc. Specifications of the sockets are generally two pin 5A/240V, three pin 5A/240V, 15A/240V, two-in-one 5A/240V, 15A/240V. In 3 or 5 pin socket line point always right side, neutral point always left side and top point always earth terminal.



Plugs

The purpose of plug is to take supply from the sockets. The sockets may either two pin or 3 pin sockets. In two pin plug there will not be earth point but in 3 pin plug earth point will be available. In 3 pin plug one pin will be bigger in size it is called earth pin. Plugs are made up of bakelite or PVC materials. The specifications of the plugs are normally two pin 5A/240V, 3 pin 5A/240V, 3 pin 15A/240V.

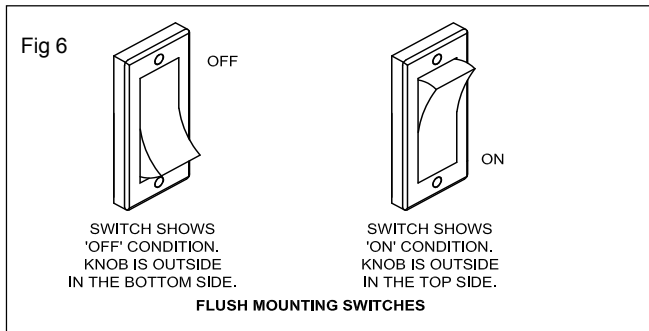


switches

In electrical engineering, a switch is an electrical component that can break an electrical circuit, interrupting the current or diverting it from one conductor to another.

The most familiar form of switch is a manually operated electromechanical device with one or more sets of electrical contacts, which are connected to external circuits. Each set of contacts can be in one of two states: either "closed" meaning the contacts are touching and electricity can flow between them, or "open", meaning the contacts are separated and the switch is non conducting. The mechanism actuating the transition between these two states (open or closed) can be either a "toggle" (flip switch for continuous "on" or "off") or "momentary" (push-for "on" or push-for "off") type.

An ideal switch would have no voltage drop when closed, and would have no limits on voltage or current rating



Different Types of Switches

Toggle Switch

Toggle switches are triggered by a device in one of two or more positions. Household light switch wiring is an example of a toggle switch. The toggle switch is usually a two position switch. The inside structure has a spring such that it is held tightly in the position to which it has been switched. It is generally used for switching purposes on or off. Many versions are able to tolerate switching 250 VAC at current levels of around 1 amp.



Push Button Switch

Push button switches are mostly used in many electronic circuits and design. These switches required to give a momentary connection. Instead they can be used to deliver a push on - push off action. Pushbutton switches are two-position devices triggered with a button that is pushed and released. Most of it has an internal spring tool that returns the button to its "out" position and "in" position.

Push Button Switch Positions



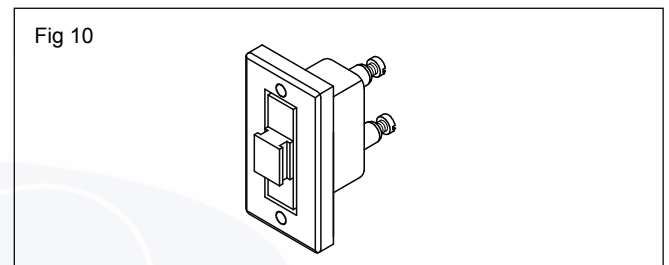
Push Button Switch

Rotary Switch

As the name denotes, rotary switches are activated by rotating a knob. Choosing the correct position allows the relevant connections to be completed. As rotary switches can have different positions, they permit a particular point to be connected to one of a number of other points in the electronics circuit. Another type of switch is rotary switch consists of three levels attached on a common shaft.

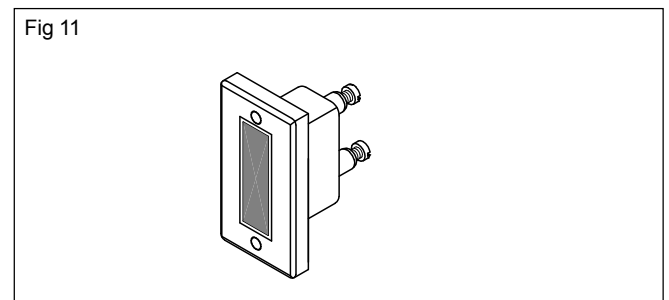
Fuses

Fuse is meant for break the electric supply when short circuit or over current takes place and it will act as a protective device. There are several types of fuses are available in the market, they are flush type, glass fuse, kit-kat fuse, HRC fuse and semiconductor fuse etc. In fuses slow blow type and fast blow type according to the equipment/device are also available in the market. The specifications of the fuses are normally 100mA, 200mA, 500mA (glass type), 5A/240V (flush type) 25A, 50A, 75A, 100A HRC (High rupturing capacity)/Semiconductor type.



Indicators

The purpose of providing indicator in equipment/device is to identify the electric supply is available or not, after switch on the main supply. Built in type with switch and separate neon indicator lamps are generally used. We can find indicator in all electrical as well as electronics equipments such as non auto/auto iron box, xerox machine, heaters, power supplies and computers etc



The specification of the indicators are 5A/240V flush type, 12V/24V indicators lamps etc.

Space for recording additional information



Meter range & measure types

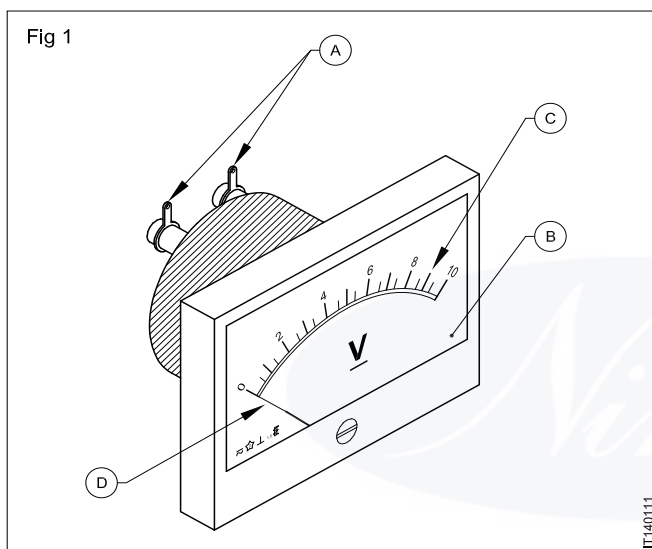
Objectives: At the end of this lesson you shall be

- state the use of meters
- list the basic parts of a simple meter
- list the minimum specifications that go with any meter
- list the symbols used on meter dial and interpret their meaning.

Meters are instruments used to measure electrical quantities like voltage, current, resistance etc.,

Measurement of electrical quantities is necessary while installing, operating, testing & repairing electrical & electronic equipments and circuits.

A simple meter is shown in Fig 1.



The electrical quantity to be measured is given to the input terminals (A) of the meter. The internal meter movement or mechanism moves the pointer (D) over the graduated scale (C) marked on a plate called the dial plate (B). The pointer stops at a point on the scale which corresponds to the magnitude of the input given at the input terminals (A).

Any simple meter must have the following minimum specifications.

- [1] The electrical parameter it can measure.

Example: DC voltage, AC voltage, DC current, AC current, resistance and so on.

- [2] The maximum quantity that it can measure.

Example: 10 volts, 100 volts, 1 ampere and so on.

The simple meter shown in Fig 1 can measure DC voltage. This can be found out from the symbol $\underline{V}$ marked on dial plate of the meter. All meters will have such symbols by which the user can identify the electrical parameter that the meter can measure. The different symbols used and their meanings are shown in Charts 1 at the end of this lesson.

Example 1: A symbol $\underline{V}$ on a meter dial indicates,

$\underline{V}$ for measuring voltage

$\sim$ for measuring AC.

$\sim$

This means, a meter with $\underline{V}$ symbol is for measuring AC voltage.

$\sim$

Example 2: A symbol $\underline{V}$ on the meter dial indicates,

$\underline{V}$ for measuring voltage

$\sim$ for measuring AC

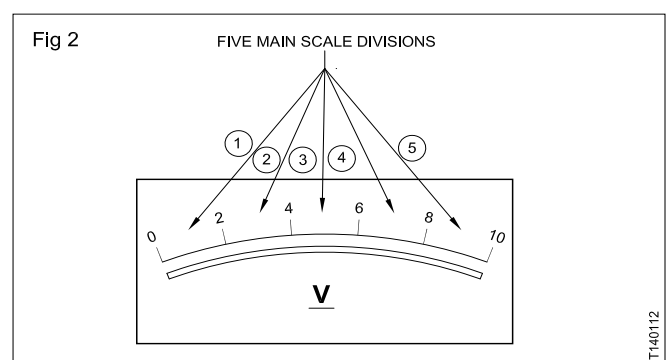
$\underline{\quad}$ for measuring DC.

$\sim$

This means, a meter with $\underline{V}$ symbol is for measuring AC and DC voltages.

The meter scale as shown in Fig 1, is graduated/marked from 0 to 10. This means that this meter can measure up to a maximum of 10 volts. This is referred to as the maximum measurable value in that meter.

The meter scale of 0 to 10 is divided to 5 parts in steps of 2 volts as shown in Fig 2. Each division is called the **Main Scale Division (MSD)** of the meter scale.

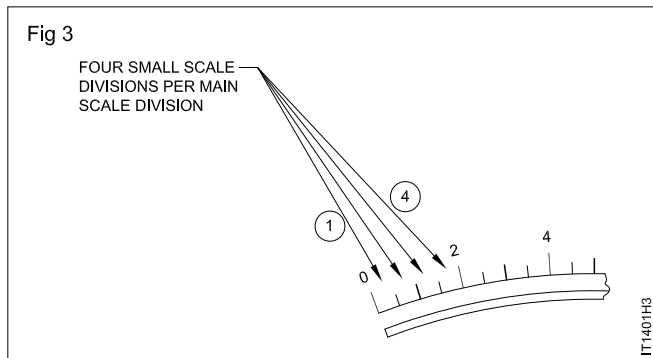


Each main scale division in Fig 2 corresponds to 2 volts. Further each main scale division (say 0 to 2) is further divided into 4 more divisions as shown in Fig 3. These divisions are called **Small Scale Divisions (SSD)**.

Each small scale division therefore corresponds to,

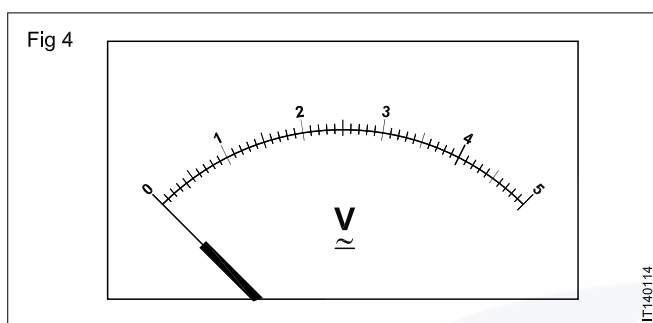
$$\frac{\text{Value of one main scale division}}{\text{Number of small scale division per main scale division}}$$

for fig. 1, each SSD is, $\frac{2 \text{ volts}}{4} = 0.5 \text{ volts}$



Hence the smallest voltage that can be accurately measured using this meter is 0.5 volts. This is nothing but the value of one small scale division of the meter.

Example: To find the maximum and minimum values that can be measured using a meter having a graduated scale as shown in Fig 4.



Maximum quantity the meter shown in Fig 4 can measure is equal to the full scale deflection value or the highest numeric on the right edge of the of scale = 5 volts.

Minimum quantity the meter can measure is equal to value of one small scale division

$$= \frac{\text{value of one main scale}}{\text{No. of small scale div/main div}}$$

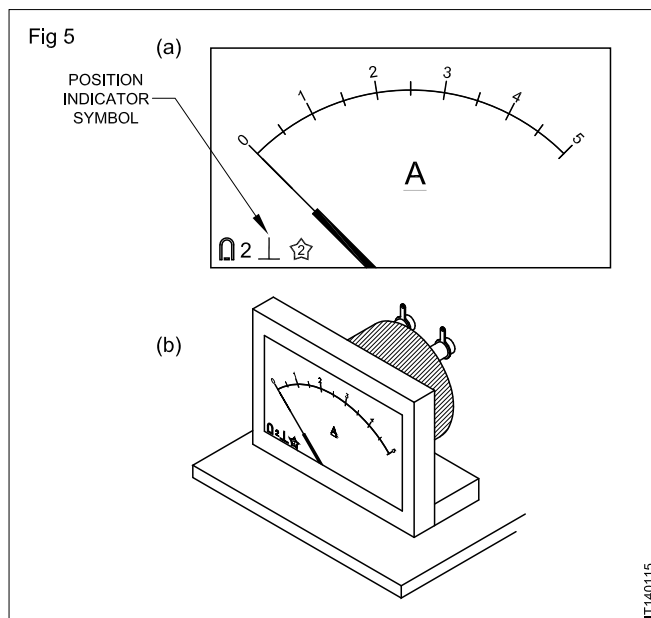
$$= \frac{1 \text{ Volt}}{10 \text{ div}} = 0.1 \text{ volts.}$$

The minimum values that can be measured using the meter in Fig 4 is 0.1 volts and the maximum values that can be measured is 5 volts.

On the dial scale of any meter, in addition to the symbols indicating the electrical parameter (voltage, current etc) it can measure and the type of parameter(ac,dc,ac/dc), there are several other symbols. One of the important symbols to be identified before using the meter is the position symbol.

Fig 5(a) indicates a typical position symbol on the dial plate of a meter.

' ⊥ ' symbol on the dial plate indicates that, the meter has to be positioned vertically (at right angle to the Table) as shown in Fig 5(b). If this meter is placed horizontally while taking measurement then, the readings shown by the meter will not be accurate.



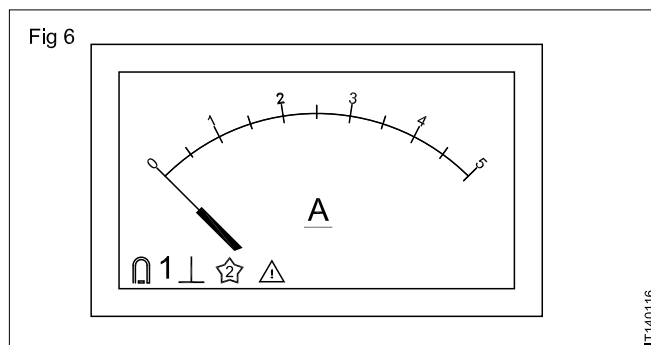
Other symbols indicating the position in which a meter is to be kept while taking readings is given in the Chart 1 of this lesson.

[H.I Use Chart 1 and elaborate the meaning of the symbols in the classroom.]

CLASSROOM EXERCISE

For a meter with dial plate markings as shown in Figure 6, referring to Chart 1 of this lesson, find;

- the nature and type of electrical parameter it can measure
- the position in which the meter is to be kept while using
- the type of mechanism used for the pointer movement
- the percentage of error indicated in the meter reading
- the maximum test voltage that can be applied
- the minimum and maximum quantity the meter can accurately measure.



The meters discussed so far, can measure only one range of values. The meter shown in Fig 1 can measure 0 to 10 volts. The meter shown in Fig 4 can measure 0 to 5 volts. Such meters are called **Single range meters**. These meters are generally mounted on electrical panels, and on the front panel of power supply units. Hence, these meters are commonly referred to as panel meters.

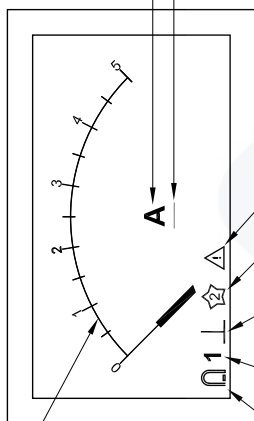
Chart 1

SYMBOLS ON METERS

Scale Marking
Linear / Non Linear

Indicates
type of meter

Indicates the
reading of AC/DC



Symbol on meter dial	Meaning of the Symbol
V	VOLTMETER
mV	milli-VOLTMETER
mV	micro-VOLTMETER
A	AMMETER
mA	milli-AMMETER
mA	micro-AMMETER
W	OHMMETER
OHMS	OHMMETER

Symbol	Meaning of the Symbol
	DC VOLTAGE OR CURRENT
	DC VOLTAGE OR CURRENT
	AC VOLTAGE OR CURRENT
	AC / DC VOLTAGE OR CURRENT

Indicates the type of mechanism/principle of the meter pointer movement associated with the meter.

Indicates percentage error in the indicated meter reading.

Indicates the placement position of the meter

Indicates special instructions that go with the meter.

Indicates special instructions that go with the meter.

Symbol	Meaning of the Symbol
	MOVING COIL
	MOVING COIL WITH PERMANENT MAGNET
	MOVING COIL WITH RECTIFIER
	MOVING COIL WITH THERMAL CONVERTER/ THERMO COUPLE
	MOVING COIL WITH THERMAL CONVERTER ISOLATED
	MOVING COIL QUOTIENT MEASURING
	INDUCTION TYPE
	INDUCTION QUOTIENT MEASURING
	HOT WIRE
	BIMETALLIC
	VIBRATING REED

Symbol	Meaning of the Symbol
	MOVING IRON
	MOVING IRON
	MOVING IRON QUOTIENT MEASURING
	MOVING COIL NEEDLE / MAGNET
	MOVING MAGNET QUOTIENT MEASURING
	ELECTRO DYNAMIC CORELESS
	ELECTRO DYNAMIC QUOTIENT MEASURING CORELESS
	ELECTRO DYNAMIC WITH IRON CORE
	ELECTRO DYNAMIC QUOTIENT MEASURING WITH IRON CORE
	ELECTRO STATIC

Symbol	Meaning of the Symbol
	VERTICAL POSITION
	VERTICAL POSITION PERMISSIBLE DEPARTURE ± 2° FROM VERTICAL POSITION
	HORIZONTAL POSITION
	HORIZONTAL POSITION PERMISSIBLE DEPARTURE ± 10° FROM HORIZONTAL POSITION
	INCLINED POSITION ANGLE OF INCLINATION 60° FROM HORIZONTAL

Symbol	Meaning of the Symbol
	NO TEST VOLTAGE
	TEST VOLTAGE 500VOLTS
	TEST VOLTAGE 1KLOVOLTS
	TEST VOLTAGE 2KLOVOLTS

Symbol	Meaning of the Symbol
	MAGNETIC SHIELD
	ELECTROSTATIC SHIELD
	ATTENTION READ INSTRUCTIONS BEFORE USE..

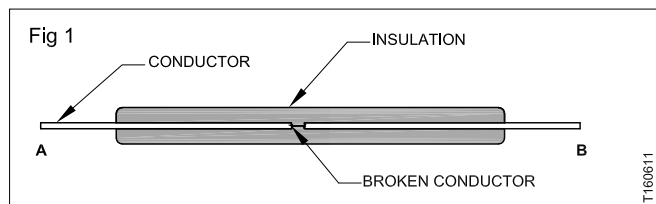
Simple AC, DC Circuits and measure current & voltage

Objectives : At the end of this lesson you shall be able to

- state the meaning and need of continuity test
- define an electric/electronic circuit
- state the method of tracing multiwire cable
- state the method of testing very lengthy pair of wires
- state the method of testing ON/OFF switches
- state the method of testing proper grounding in electrical appliances.

CONTINUITY TESTING

Due to continuous usage and/or rough handling, wires, cables, meter probes etc, becomes defective. One of the most common defect in wires and cables is that, the conductor(s) inside the insulation of the wire gets cut/ broken as shown in Fig 1.



Such defects in wires are not visually detectable because of the insulation over the conductor. The only way to check such a defect is by checking whether continuity exists along the wire or not.

A wire (conductor) without a cut will have *zero ohms of resistance* across its ends. Whereas a conductor *with a cut* will have almost *infinite resistance*. Checking whether the wire has zero resistance or infinite/very high resistance is referred to as *continuity testing*.

Referring Fig 1, due to break in the conductor, the continuity along A and B is broken. Hence, the resistance across A and B will be very high.

Such continuity checks can be carried out, by making the wire to be checked, as a part of a circuit and checking if the circuit functions or not.

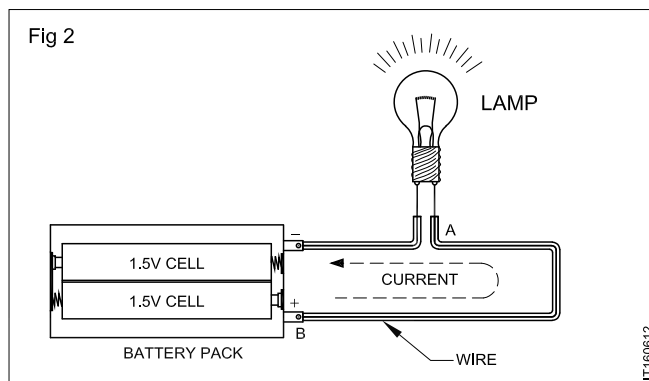
CIRCUIT

A circuit is nothing but, several electrical, electronic components connected in a pre-defined way and energised by a voltage source.

A circuit can be defined as any closed path consisting of two or more electrical/electronic components through which electric current can flow. Fig 2 shows a simple circuit consisting of a battery pack and a lamp interconnected using wires.

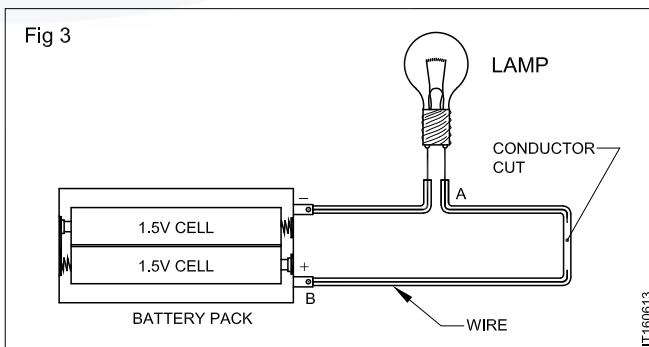
In the circuit in Fig 2 the direction of arrows indicates the direction of current flow.

Conventional current flows from +ve terminal of the battery into the circuit and enters the negative terminal of the battery.



The voltage source makes the current to flow through the closed circuit. Since, the lamp is a part of this closed circuit, the lamp glows.

If a section of wire AB in the circuit at Fig 2 is cut/broken at some point as shown in Fig 3, then, the closed path of the circuit is broken. Hence, current can no more flow through the circuit and the lamp do not glow.



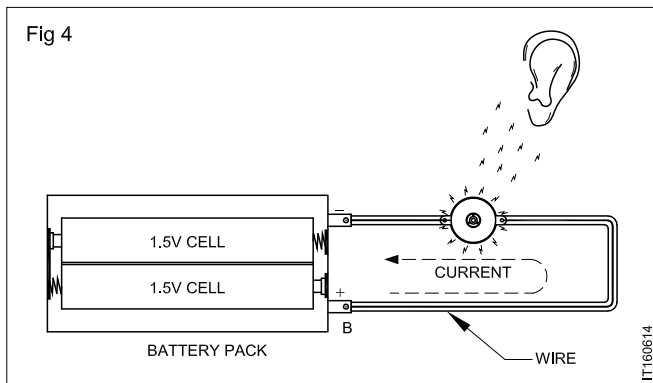
From the above it is evident that,

- the lamp *glows* if wire AB has continuity or does not have **breaks or cut**
- the lamp *does not glow* if wire AB has a **discontinuity** or breaks or cuts.

Hence, this circuit shown in Fig 2, can be used as a *continuity tester* by connecting a wire to be tested across A and B of the circuit. This circuit can, therefore, be called a **continuity tester**.

Instead of a lamp, a buzzer can be connected to have the aural indication as shown in Fig 4.

A buzzer is an electromechanical device which produces a buzzing sound when current is made to flow through it.



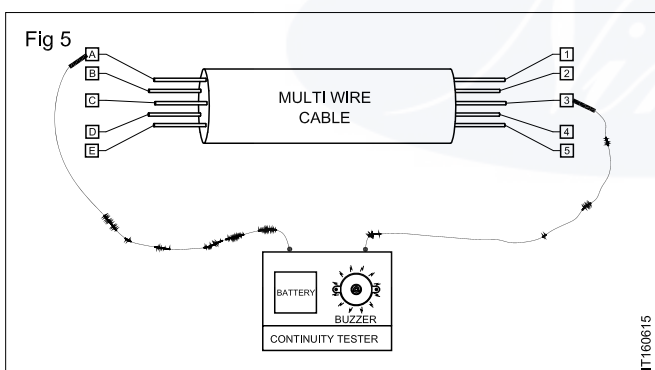
Buzzers are available for various voltage ratings, starting from 1.5V to several hundred volts. Buzzers are available for both AC and DC supplies.

Continuity tests although can be carried out using an ohmmeter, simple aural or visual continuity testers are easy and is cost-effective. Hence, continuity tester is one of the very important test instruments for electrical and electronic technicians.

In addition to testing the continuity of wires, continuity testers find several applications as given below;

i) Tracing wiring of multicore cables

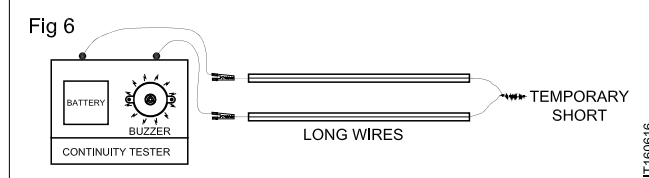
When several wires in a bunch run between any two sets of points as shown in Fig 5, continuity testers can be used to trace which of the several wires runs between any two points.



This is done by checking the continuity of A with end points 1, 2, 3, 4, 5. The point at which the continuity tester gives a buzzing sound will be the point connected to A. In the example given in Fig 5, points A and 3 belongs to the same wire.

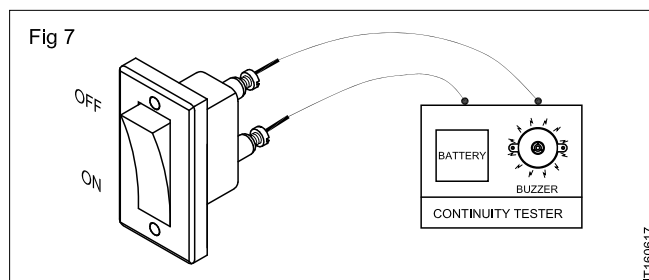
ii) Continuity of two wire lines and cables

When it is required to check the continuity of two wire lines whose ends are too far apart as shown in Fig 6, any one end of the two conductors can be temporarily shorted and the continuity of both the wires checked at the other end.



iii) Testing working condition of electrical switches

As shown in Fig 7, for a switch to be declared good, the tester must buzz when the switch is put to the on position and should not buzz when put to the off position.

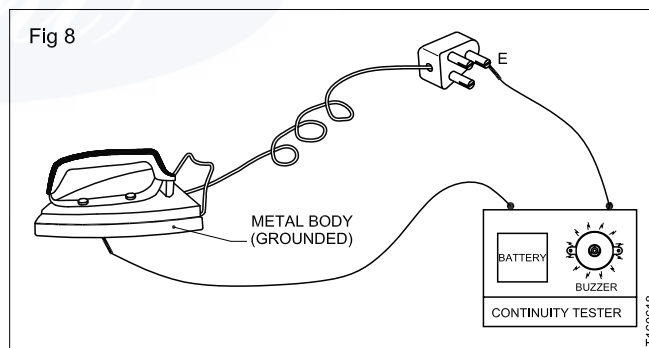


Almost all types of electrical switches can be tested using a continuity tester.

iv) Testing proper grounding in small electrical equipments/ appliances

If the ground wire is properly connected, the continuity tester must buzz when connected as shown in Fig 8.

Continuity testers can also be used to check if there is *short* between the body of the iron and *live*, *neutral* points of the mains plug.



Multimeter

Objectives: At the end of this lesson you shall be

- identify multimeters - different types - functions of the main parts, controls-schematic diagrams showing internal connections for different modes of operation
- explain the method of measuring direct/alternating voltages with a multimeter
- explain the method of measuring direct current with a multimeter
- explain the method of measuring resistance by a multimeter
- explain the precautions to be observed while measuring voltage, current and resistance in the circuit
- state the applications of a multiplier
- state the types of multimeter
- explain the working of digital multimeter
- compare analog and digital multimeter.

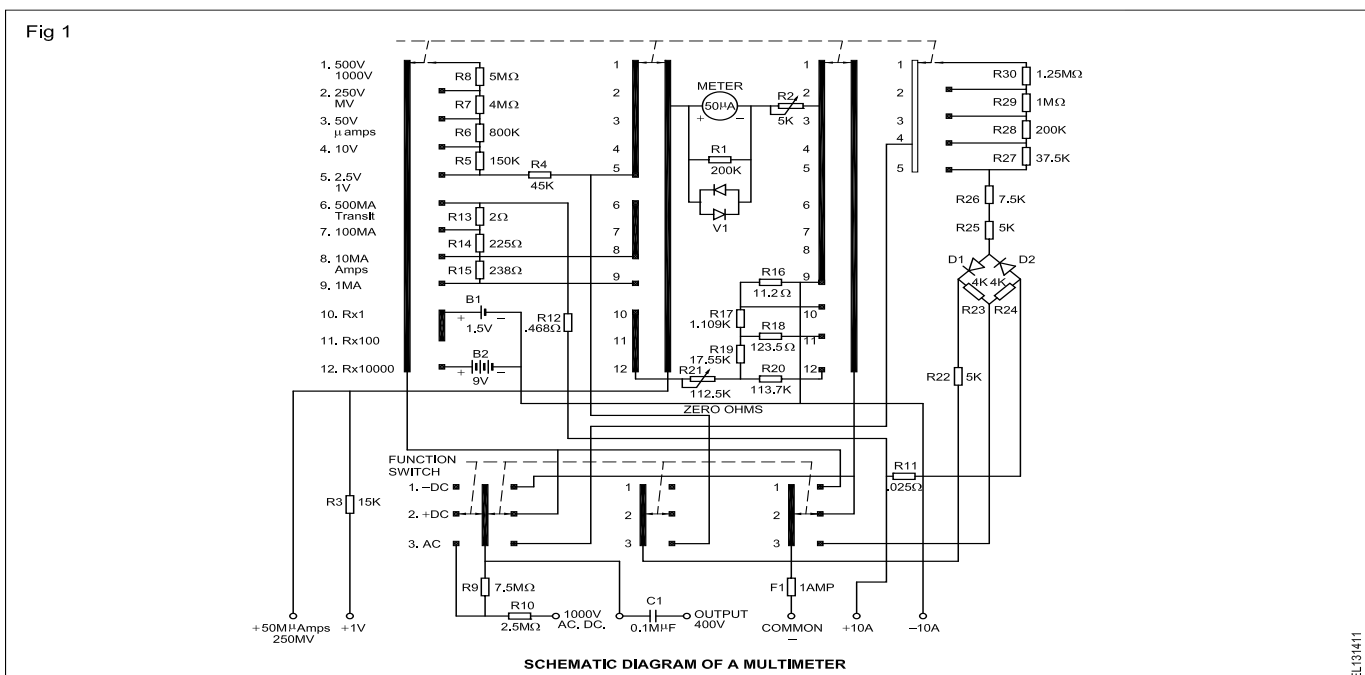
Multimeter: A multimeter is an instrument in which the function of an ammeter, voltmeter and ohmmeter are all incorporated for measurement of current, voltage and resistance. Some manufacturers call this as VOM meter as this meter is used as a volt, ohm and milliammeter. The multimeter uses the basic d'Arsonval (P.M.M.C) movement for all these measurements. This meter has facilities through various switches to change the internal circuit to convert the meter into a voltmeter, ammeter or ohmmeter. There are several types of multimeters available in the market, manufactured by various manufacturers. Each model differs from the others by the extra facilities available. It is a versatile tool for all electricians. With proper usage and care, it could give service for many years.

Types of multimeters: There are two types of multimeters available in the market. They are:

- ordinary multimeter having passive components in the internal circuitry
- electronic multimeter having active and passive components in the internal circuitry.

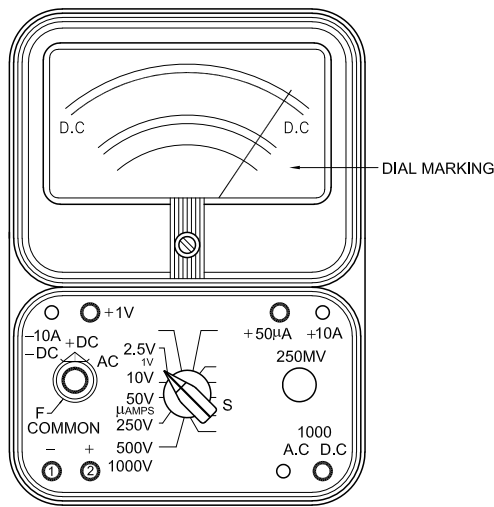
Though the ordinary multimeters are of an analog type, the electronic multimeters can be either analog or digital type. Most of the ordinary multimeters will have a sensitivity of 20 K ohms per volt whereas electronic multimeters have internal resistance to the tune of 5 to 10 megohms, irrespective of the selected ranges. The explanation given here is for an ordinary multimeter only.

Fig 1 shows the schematic diagram of a commercial multimeter. This meter is a combination of a multi-range DC voltmeter, AC voltmeter, DC milliammeter, an ohmmeter and an output meter. The indicating meter in this multimeter uses the d'Arsonval movement which has a resistance of 2000 ohms, having full scale deflection current of 50 micro-amperes.



This instrument is provided with a function switch 'F' to select the type of supply, whether AC or DC. Further the same switch can be used to select the polarity of the meter in DC mode without changing the leads as indicated by position +DC or -DC in Fig 2.

Fig 2

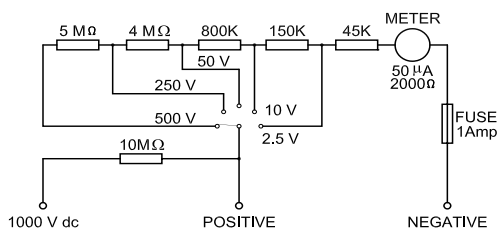


A single range switch S at the centre selects the range of AC or DC voltages, DC current (no facility to measure alternating current in this meter) and resistance.

The range switch, along with the function switch, connects the required circuit elements to convert the meter for the required measurement of electrical quantities as described below.

Multimeter as a DC voltmeter: The circuit shown in Fig 3 is for DC voltmeter section when common input terminals are used for voltage. The voltage range can be selected with the help of the range switch. The circuit in Fig 3 shows the value of the multiplier resistance for various ranges at DC

Fig 3

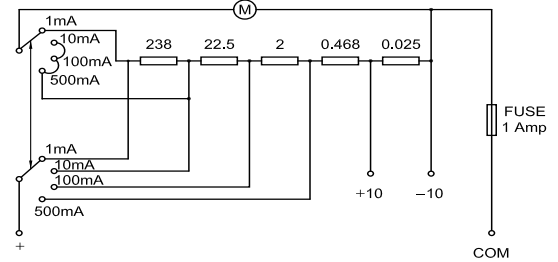


SCHEMATIC DIAGRAM OF MULTIMETER AS DC VOLTMETER

voltage mode. This instrument has a sensitivity of 20K ohms/volt in the DC mode which is fairly high. Hence suitable for electronic measurements also. A separate jack is provided for measuring voltage 1000V. Apart from using separate jacks, the range switch should be set to 500V position for measuring 1000V.

Multimeter as a DC ammeter: Fig 4 shows the circuit configuration of this multimeter for measuring DC milliampere or ampere. For measuring DC current up to 500mA common jacks are used whereas the jacks marked +10 and -10A should be used for 0-10A range.

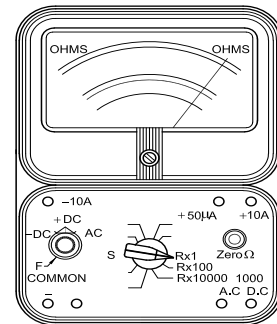
Fig 4



SCHEMATIC DIAGRAM OF THE MULTIMETER IN DC CURRENT MODE

Multimeter as an ohmmeter: For the ohmmeter function, the range switch should be set as shown in Fig 5 and the

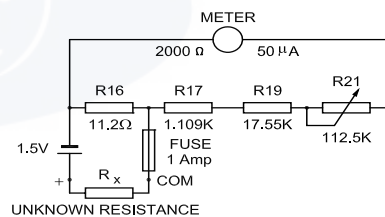
Fig 5



function switch should be in DC mode. The details of the ohmmeter section of this multimeter are shown in Fig 6.

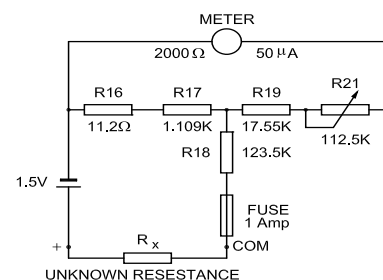
Fig 6a shows the ohmmeter section for a scale multiplication of x1 range, 6b for x100 and 6c for x10000 ohms ranges.

Fig 6(a)



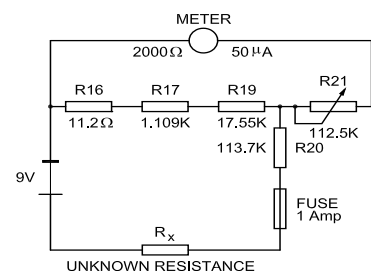
RESISTANCE MODE Rx1 RANGE

Fig 6(b)



RESISTANCE MODE Rx100 RANGE

Fig 6(c)



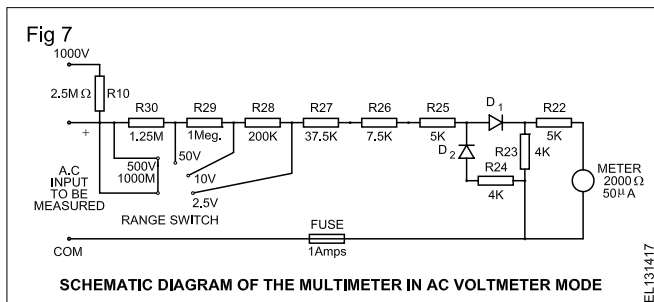
RESISTANCE MODE Rx10k RANGE
SCHEMATIC DIAGRAM OF THE MULTIMETER IN OHMMETER MODE

Though resistance ranges $\times 1$ and $\times 100$ uses a single cell, the $R \times 10000$ range uses an additional battery of 9V.

Note that the meter supply polarities to the unknown resistance is reversed in the multimeter. This hint will be useful when testing diodes and transistors.

Multimeter as a AC voltmeter: For the AC voltage measurement the function switch should be in AC position and the range switch should be set to the suitable range.

Fig 7 shows the circuit arrangement for measuring AC voltages. As shown in figure 7, D₁ is used to convert AC input to DC for meter movement. Due to the use of a rectifier, the sensitivity in an AC voltmeter mode for this multimeter is only 5000 ohms/volts..



The method of using a multimeter for measuring various quantities is as follows.

Methods of measuring DC voltages: Refer to Fig 2, set the function switch (F) in the DC position and set the range switch (S) as per the voltage to be measured, plug in the black lead to the jack(1) marked 'Common', the red lead to the +ve jack(2). Connect the test prods to the voltage source to be measured.

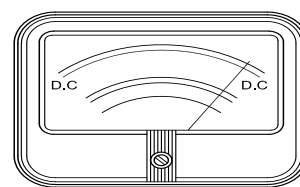
When measuring DC voltage, care must be taken to prevent damage to the instrument by applying a higher voltage than the voltage indicated in the range setting. For this, while measuring voltages, it is always better to select the highest range first and then reduce to the lower range to have either a full scale deflection or a reading which is above mid- scale.

For example, if the percentage error is 3% for the full scale deflection, at the middle of the scale the error goes to 6% and for 1/4 scale the deflection error shoots up to 12%. Hence it is advisable to get the reading somewhere between the centre of the scale and the full scale.

If higher voltages are measured in low range settings, the meter will get damaged.

On the other hand if low voltages are measured in higher range settings, the deflection will be less and the percentage error will be more. Hence, a proper range is to be selected to have a deflection more than half of the dial. Refer to Fig 8.

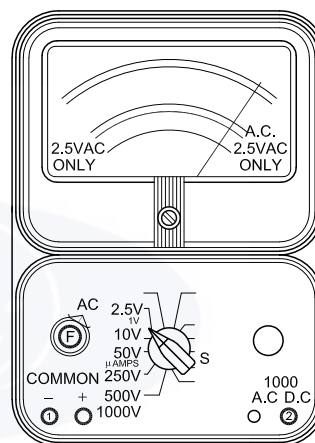
Fig 8



Method of measuring AC voltages: The function switch should be set at AC position, and the range switch to a proper voltage position. Plug the black test lead in the COMMON jack and the red lead in the +ve jack. Then connect the test leads across the voltage source. For 0-2.5V range, read the value directly on the scale marked 2.5V AC only.

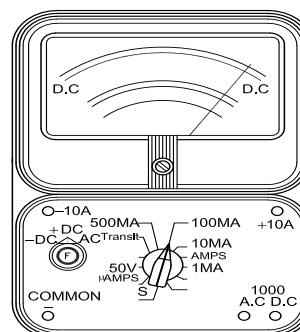
Refer to Fig 9 for 10V, 50V and 250V ranges, read the red scale marked AC and use the red figure below the scale. For the 500V range read the red scale marked AC having 0-50 scale and multiply the reading by 10.

Fig 9



Method of measuring direct current: Always connect the instrument in series with the load. Set the function switch in DC position and the range switch on the desired range of current or preferably at a higher range. The black lead is to be connected to the 'COMMON JACK' and the Red lead to the +ve jack. The current in milliamps should be read on the black scale. Refer to Fig 10.

Fig 10



In all direct current measurements, be certain that the power to the circuit being tested has been turned 'OFF' before disconnecting or connecting the test leads. For 10 amps current measurement, use jacks marked +10 and -10 only.

When using the instrument in 10A range never remove the test lead from its panel jack while the current is flowing through the circuit. This causes damage to the banana plugs and jacks.

Method of measuring resistance: When resistance is to be measured, set the function switch 'R' to + DC position and set the range switch 'S' to the resistance position of the desired range as shown in Fig 5. Plug the black lead in the common jack and the red test lead in the jack. Short the lead and check the zero position of the pointer. If the pointer is not in zero position, adjust the zero ohms adjustment knob until the pointer sets on the zero position. If the pointer does not come to the zero position of the ohmmeter, the battery should be replaced.

Before measuring resistance in any circuit, be sure the power is off in the circuit being tested. Discharge any electrolytic capacitor in the circuit before measuring the resistance.

Method of use:

Use R x 1 for resistance values from 0 to 200 ohms.

Use R x 100 for resistance values from 200 to 20,000 ohms.

Use R x 10,000 for resistance values for above 20,000 ohms.

The topmost scale on the dial is intended for resistance measurement.

Application of a multimeter: A multimeter is used for testing and fault finding in electrical/electronic circuits, electrical appliances and machines. An understanding of its proper use helps in achieving better results.

It is a handy instrument electricians use for:

- checking continuity of circuits, appliances and devices
- measuring/checking the supply presence at the source
- checking the components for their condition by resistance measurement for voltage drop across it in on line measurement
- measuring the current drawn by the circuit to infer its condition
- measuring resistance of the electrical appliances and devices for knowing their condition.

Other applications of the multimeter are for testing capacitors, diodes and transistors to know their condition. Some meters have provision also for temperature measurement with suitable sensing prods.

Types of multimeters

Analogue types:

- Selector switch type
- Multi-plug type

Digital types:

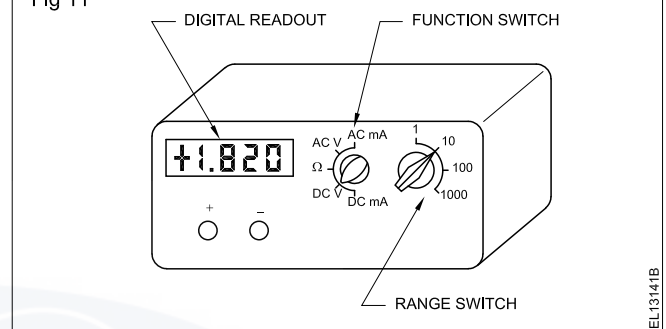
- Selector switch type
- Auto-ranging type
- LCD display

Analogue type multimeters: They have a meter movement mechanism, a calibrated scale and a pointer. Reading is obtained by looking at the position of the pointer on the scale.

The multimeter explained earlier is an analogue type.

Digital multimeter: In a digital multimeter, the meter movement is replaced by a digital readout. (Fig 11) This readout is similar to that used in electronic calculators. The internal circuitry of the digital multimeter is made up of digital, integrated circuits. Like the analog-type multimeter, the digital multimeter has a front panel switching arrangement. The quantity measured is displayed in the form of a four digit number, with a properly placed decimal point. When DC quantities are measured, the polarity is identified by means of a + or – sign displayed to the left of the number.

Fig 11



Comparison of analogue and digital multimeters

Analogue type	Digital type
1 Instrument has moving exercised. Position of use fixed.	There are no moving pointer. Care should be parts. Used in any position.
2 Reading error due to parallax is possible.	No reading error as they have numerical display.
3 Actual value of indication obtained by computation.	No need to compute. Direct indication of value.
4 Manual zero setting for resistance measurement is required.	Automatic zero setting for resistance measurement.
5 Auto-range setting not possible.	Auto-ranging instruments are available.
6 It can track short term variations and trends in measured quantity.	Not possible - slow response.
7 Indication of measured quantity by the pointer over the graduated scale.	Digital numerical read-out.
8 Loads the measuring of circuit.	Practically no loading.

DMM functions: The basic functions found on most DMMs are the same as those on analogue multimeters. That is:

- Ohms
- DC voltage and current
- AC voltage and current.

Some DMMs provide special functions such as transistor or diode tests, power measurement, and decibel measurement for audio amplifier tests.

DMM displays: DMMs are available with either LCD (liquid-crystal display) or LED (light-emitting diode) read-outs. The LCD is the most commonly used read-out in battery-powered instruments because it requires only very small amount of current. A typical battery-powered DMM with an LCD read-out operates on a 9 V battery that will last from a few hundred hours to 2000 hours and more. The disadvantages of LCD read-outs are that (a) they are difficult or impossible to see in poor light conditions, and (b) they are relatively slow to respond to measurement changes. LEDs, on the other hand, can be seen in the dark, and respond quickly to changes in measured values. LED displays require much more current than LCDs, and, therefore, battery life is shortened when they are used in portable equipment.

Both LCD and LED-DMM displays are in a seven-segment format as shown in Fig 12.



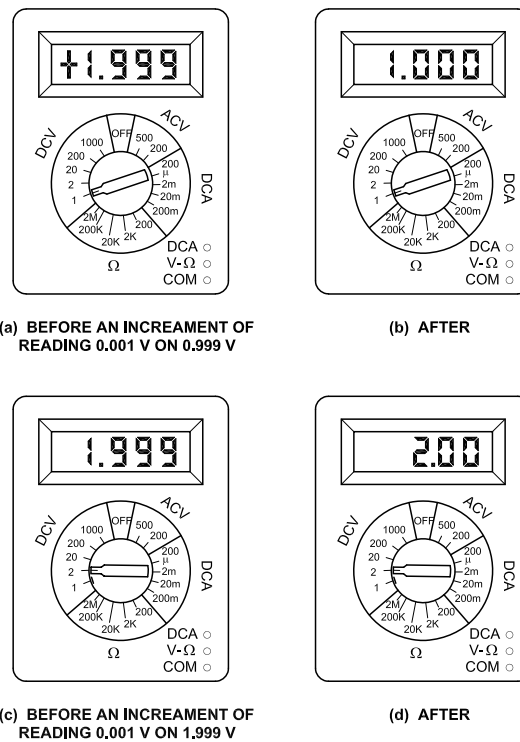
Many meters have 3 digits in their display. We will use this case for illustration. A 3 digit multimeter has three digit positions that can indicate from 0 through 9, and one digit position that can indicate only a value of 1. This latter digit, called the half-digit, is always the most significant digit in the display.

For example, suppose a DMM is reading 0.999 volt, as shown in Fig 13a. If the voltage increases by 0.001 V to 1V, the display correctly shows 1.000 V, as shown in Fig 13b. The 1 is the half-digit. Thus, with 3 digits, a variation of 0.001 V, can be observed.

A 3 digit DMM illustrates how the number of digits in use changes.

Now, suppose the voltage increases to 1.999 V, this value is indicated on the meter as shown in Fig 13c. If the voltage increases by 0.001V to 2V, the half-digit cannot display the '2', so the display shows 2.00. The half-digit is blanked and only three digits are active, as indicated in Fig 13d. DMMs with displays of 4 through 8 digits are also available.

Fig 13



Multimeter - Safety precautions: Using a multimeter safely is an important technical skill. When measuring electricity, you are dealing with an invisible and often a lethal force. Voltage levels above 30V can kill. The following safety precautions should always be followed.

- Never use the ohmmeter section on a live circuit.
- Never connect the ammeter section in parallel with a voltage source.
- Never overload the ammeter or voltmeter sections by attempting to measure currents or voltages far in excess of the range switch setting.
- Check the meter test leads for frayed or broken insulation before working with them.
- Avoid touching the bare metal clips or tips of the test probes.
- Whenever possible, remove the voltage before connecting the meter test leads into the circuit.
- When connecting the meter test leads to live circuits, work with one hand hanging at your side to lessen the danger of accidental shock.
- To lessen the danger of accidental shock, disconnect the meter test leads immediately after the test is completed.

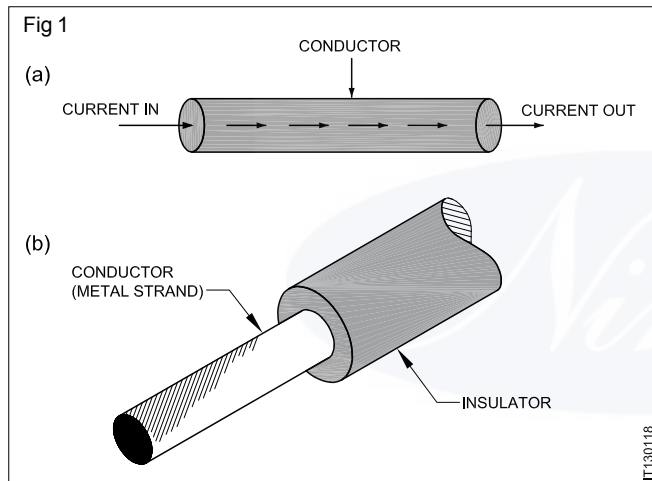
Skinning and tinning of wires

Objectives: At the end of this lesson you shall be

- define a wire and distinguish between single strand and multistrand wires
- define skinning
- state the meaning of soldering, list two main types and uses of soldering
- state the need of soldering in electrical/electronic circuits
- state the meaning of solder and types
- explain the purpose of using flux while soldering
- state the meaning of tinning
- explain the need for inspecting a tinned/soldered point.

Wires

Any metallic conductor drawn (a process of pulling) in the form of a thin solid cylindrical string, is called an electrical wire. Wires are used as a medium for carrying current from one point to another point in an electrical/electronic circuit as shown in Fig 1a.



All wires have at least one metal strand which is able to conduct electricity. This is known as the conductor. This conductor is generally surrounded by a material which does not conduct electricity as shown in Fig 1b.

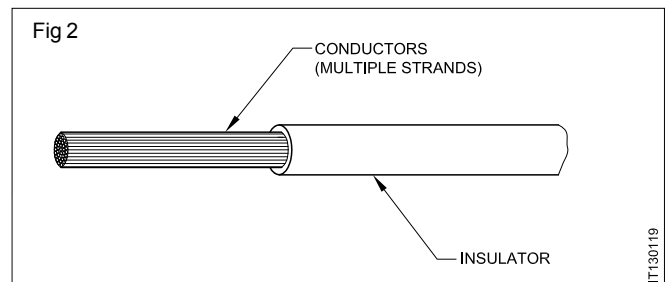
Conductors used in wires are invariably made of good to very good conductor materials like aluminum, copper. This is because electrical wires are intended to carry current with minimum opposition or resistance.

Single Strand Wires

If the conductor of a wire is a single strand of metal, as shown in Fig 8b, then the wire is known as a **single strand wire**, commonly called **hook-up wire**. These wires are commonly used for jumpering, experimental circuit wiring of electronic circuits.

Multistrand wires

If the wire consists of several strands of metal as shown in Fig 2, then the wire is known as **Multi-strand wire**.



Some of the advantages of multistrand wires over single strand wire are listed below;

- 1 More flexible and hence can be routed in any direction more easily.
- 2 More rugged and hence are less likely to cut when bend.
- 3 Can carry more current when compared to a single strand conductor of same dimension.
- 4 More suitable for high frequency application due to reduce skin effect (discussed later)
- 5 Better cooling due to air gap between strands.
- 6 Cancellation of the effect of magnetic field.

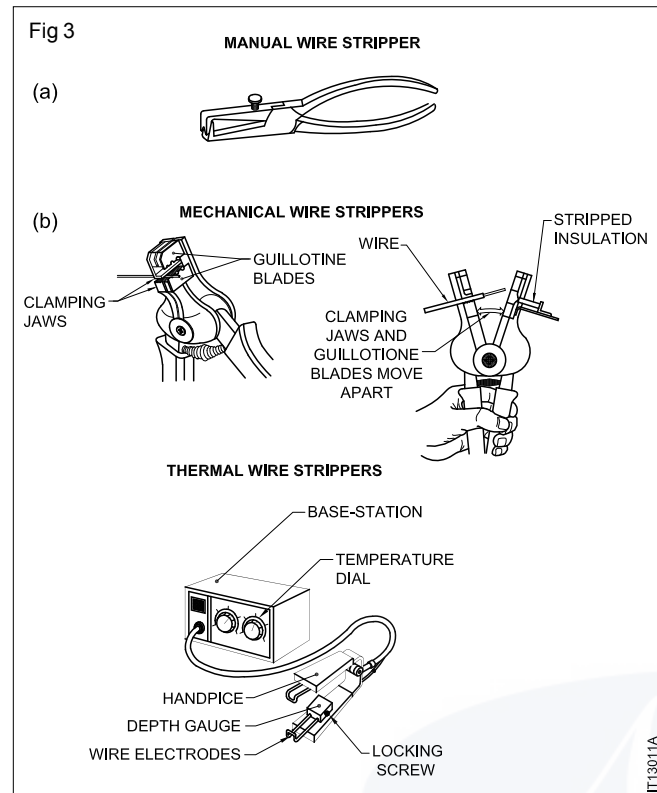
Because of the above advantages, multistrand wires are used extensively in electronic circuit connections.

Skinning of wire

To use wires in constructing electrical or electronic circuits, a small portion of insulation at the ends of the wire is to be removed. The process of removing a required length of wire end insulation is called **SKINNING** of wires. There are several methods and tools to do skinning. The simplest of all these methods is skinning using a knife.

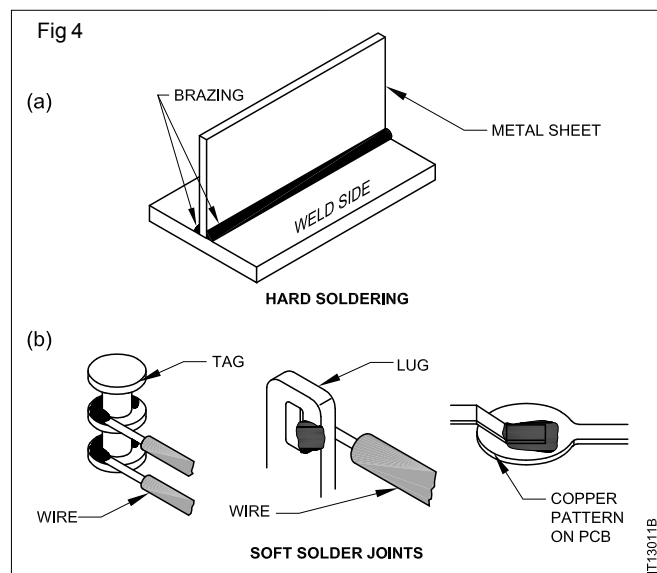
Skinning of wires using a knife is most suitable for single strand wires or multi strand wires with thick strands and thick insulation. This method is not advised to be used for skinning multi-strand wires with thin strands as the strands may get cut while skinning.

Special skinning tools are available for skinning flexible multi-stand wires. The most common of them is the manual wire stripper as shown in Fig 3a. A manual wire stripper is the cheapest of the stripping tools and does good stripping work if one practices with it. A few other types of wire strippers are shown in Fig 3b.



SOLDERING

Soldering is a process of connecting any two metallic surfaces such as copper, brass and alloys of these metals. Some types of solder joints are shown in Fig 4.



There are two types of soldering,

- [1] **Hard soldering or brazing** used for joining large metal parts as shown in Fig 4a.
- [2] **Soft soldering** used to form good electrical joints/connections between electrical/electronic parts as shown in Fig 11b.

Soft soldering is used extensively for electronic circuit wiring. In this lesson only soft soldering is discussed. Hard soldering or brazing is out of scope of this lesson.

From now on in this book, soldering means soft soldering.

Need for soldering

Requirements of an electrical joint

- [1] The electrical joint must provide ideally zero resistance or at least a very low resistance path, for the flow of current.
- [2] The electrical joint made should be strong enough to withstand vibrations, physical shock, bumps etc, without causing any deterioration to the quality and strength of the joint.
- [3] The electrical joint should be able to withstand corrosion and oxidation due to adverse atmospheric conditions.

All the above requirements of an electrical joint can be achieved by making a solder joints.

Solder

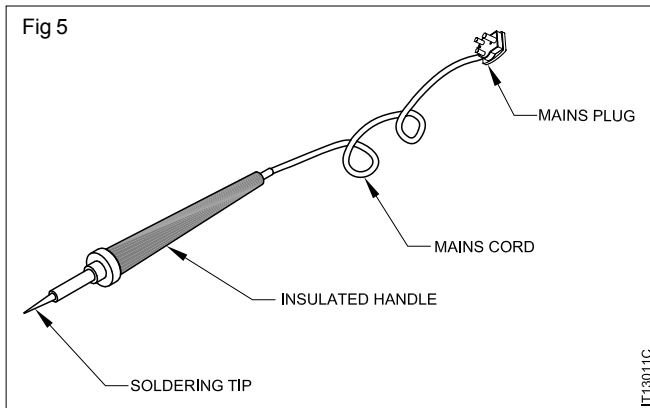
In a soldered joint, the solder is a mixture of metals, generally TIN and LEAD. It is made to melt at a certain temperature. It acts as a filler between the parts of the connection/joint to form a continuous, low resistance metallic path for conduction of electricity.

In soldering, as the metal surface is wetted (free flow of liquid solder over a surface) by the solder, a complex chemical reaction, bonds the solder to the metal surface

The tin content of the solder diffuses with the metal surface to form a layer of a completely new alloy. The alloy so formed will have the same structure as the constituent metals and retain their metallic properties and strength.

Soldering and soldering irons

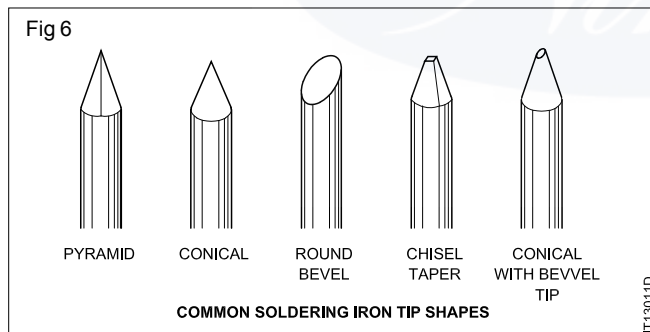
While soldering, the solder is made to melt between the metallic surfaces of the joint, using a soldering iron, as shown in Fig 5. A **soldering iron** is an instrument used to produce the required heat to carry out soldering.



Soldering irons of different wattage ratings starting from 10 watts to more than 150 watts are available commercially. Depending on the type, size and heat sensitivity of the components being soldered, the most suitable wattage soldering iron should be chosen. Most of these soldering iron work on 240V, 50Hz ac mains supply. There are special type irons which work on dc supply also. For soldering delicate components, soldering irons with temperature controlling facility are used. These are known as soldering stations.

Soldering iron tips

Soldering irons are designed to take, a variety of tip sizes and shapes as shown in Fig 6. The choice of the iron and the tip to use depends on, the nature of the joint to be soldered. A proper selection of the soldering iron and tip is important for obtaining good quality soldered joint. To solder effectively, the tip of the soldering iron must be kept clean all times.

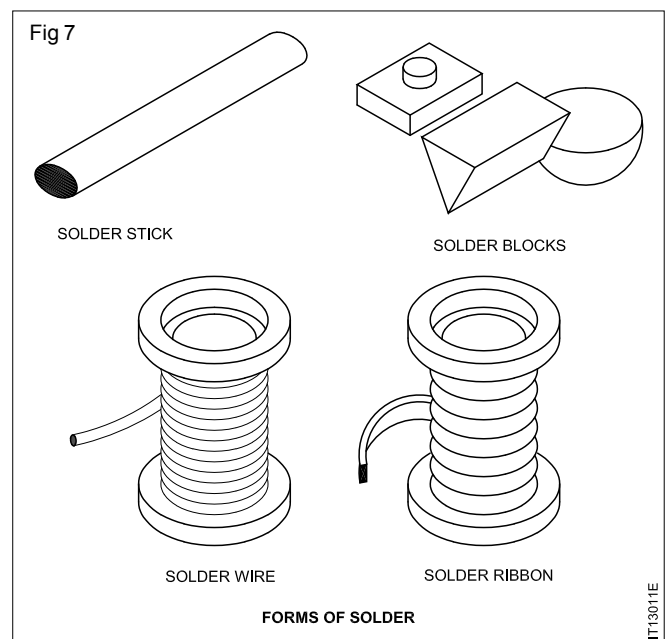


Types of solders

Solders are available in many forms as shown in Fig 7. The type to be chosen depends on, the type of soldering to be carried out.

The wire type solder shown in Fig 14 is the most commonly used solder for hand soldering work, using low wattage soldering iron.

Solders available in the market may have different tin-lead proportion in it. For general electronic circuit soldering work, solder with 60% tin and 40% lead is most suited. This solder is commonly called 60/40 solder. This solder has been specially developed to possess superior properties required for electronic circuit work.



Soldering FLUX

A protective oxide layer forms on the exposed surface of most metals. The rate at which the oxide layer is formed varies from metal to metal. The layer forms quickly on newly exposed metal, and over time, the layer slowly become quite thick.

This oxide layer on metals interferes with soldering. Hence, it must be removed before a soldered joint can be made.

The purpose of flux is to first dissolve the thin layer of oxide from the surface of the metals to be joined, and then form a protective blanket over them until the solder can flow over the joint surfaces to form the joint.

However, thick layers of oxide must be removed using an abrasive method as all types of flux are not capable of dissolving their oxide layers.

Types of flux

There are several types of fluxes used in different types of soldering. The type of flux used for soldering electronic components is called **rosin**. Rosin is made from a resin obtained from the sap of trees.

Rosin flux is ideal for soldering electronic components because, it become active at the soldering temperature, but revert to an inactive state when cooled again. An additional advantage is that it is non-conductive.

The rosin has activators or halides added to it. The activators used in rosins are mild acids that become very active at soldering temperatures. These acids dissolve the oxide layer on the metals to be soldered.

Organic and inorganic acid fluxes are available. These fluxes are not suitable for soldering electronic circuits.

For further details on the different types of rosin fluxes refer reference books listed at the end of this book.

Common forms of flux

Flux is available in a variety of forms to suit various types of application. Flux is available as a liquid, paste or a solid block. For most applications flux is often put in the solder itself during manufacture.

Not all flux types are available in all forms. For hand soldering work on electronic circuits, the best form for the flux is either as a liquid or a paste.

Rosin cored solder

Several manufacturers produce solder wire with the flux already included in one or more cores running along its length as shown in Fig 8. This is known as **cored solder**.

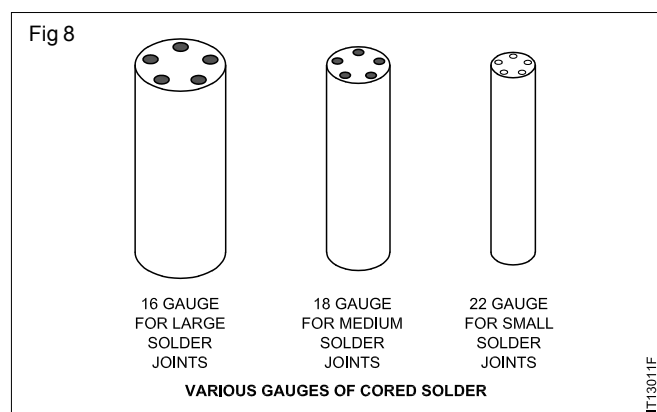
The most popular type of cored solder for electronic hand soldering contains rosin type flux. Such solder is known as **rosin cored solder**.

When the solder is heated, the rosin flux melts before the solder. The rosin then flows out over the surface to be soldered ahead of the solder.

The amount of flux contained in the core is carefully controlled by the manufacturer and for most applications it will be sufficient. However, it is a common practice to apply additional liquid flux or flux paste to the joint, just prior to making the joint. This additional flux ensures that, sufficient flux available while the joint is being made. When the soldering has been completed, excess flux if any has to be removed.

Rosin-cored solder is available in different gauges as shown in Fig 8. It is important to choose a size suitable for the job at hand as given below;

- use 22 gauge for small joints
- use 18 gauge for medium joints
- use 16 gauge for large joints.

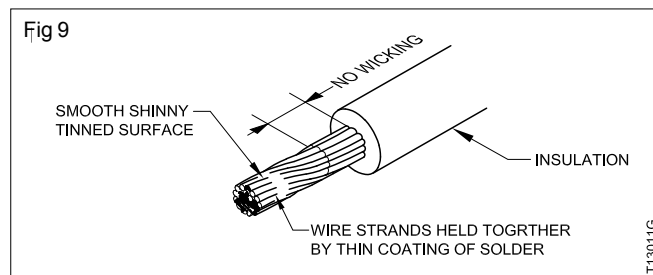


Tinning wires

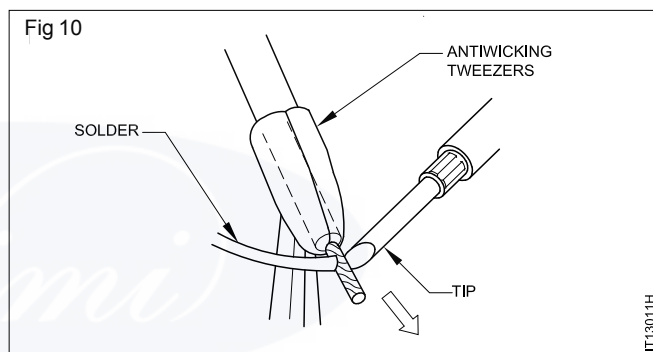
When wires are to be connected to lugs or any other type of terminations, after skinning the wire, it is preferred to apply a thin coating of solder using a soldering iron. This process is known as **tinning** the wire.

When tinned, the solder penetrates the wire strands and holds them together as shown in Fig 9. This holding of strands prevents the strands from becoming separate while soldering the wire onto terminations.

Tinning of the wire is advised to be done soon after stripping of the wire end, so that the wire strands do not tarnish.



Care must be taken when tinning a wire to ensure that capillary action does not draw the solder under the unstripped insulation. This action is called **wicking**. A special tool known as **anti-wicking tweezers** shown in Fig 10 can be used to help prevent wicking. A wire that has wicked under the insulation must be cut off and the tinning process repeated.



Inspecting tinned wires

When a wire has been tinned, it is necessary to inspect the tinned wire to ensure that:

- the solder has not wicked under the insulation
- the insulation on the wire is not melted or burnt
- the wire strands are visible beneath the solder
- the tinned surface is smooth and shiny.

If the tinned wire does not meet these standards, that portion of the wire must be cut. The wire must be re-skinned and tinned.

Soldering & De - Soldering Technique

Objectives: At the end of this lesson you shall be able to

- list the critical factors during soldering
- list the five stages and steps involved in selecting and preparation of materials
- list the different shapes of tips to be chosen for different types of soldering work
- state the functions of soldering iron stand
- state why a soldering joint should not be cooled by blowing air
- state the effect of residual flux on the joint and removing solvent
- list the features of a well soldered joint and two stages of inspection of a solder joint
- list the points to be checked during initial inspection and closer inspection
- list the common soldering defects and the effects of various types of soldering defects
- list the effects of various types of soldering defects.

Soldering a joint

Selection and preparation of the soldering materials is the most time consuming phase of making a solder joint. Heating the joint and applying solder is the least time consuming but, it is the most important part of the soldering process.

Critical factors during soldering

- 1 Controlling the temperature of the workpiece
- 2 Limiting of time that a workpiece is held at soldering temperature.

These factors are specially critical while soldering electronic components like resistors, capacitors, transistors, ICs etc., Failure to correctly time and coordinate the heating of the joint and add solder, will result in a poor quality joint and may even damage the components.

Stages in soldering

The soldering process can be divided into several distinct stages or phases as given below:

- 1 Selection and preparation of materials.
- 2 Heating the joint and adding solder.
- 3 Cooling the joint.
- 4 Cleaning the joint.
- 5 Inspecting the joint.

1 SELECTION AND PREPARATION OF MATERIALS

1.1 Selection of soldering iron wattage

Soldering irons are available in different wattage ratings starting from 10 watts to several 100 watts. The wattage of a soldering iron specifies the amount of heat it can produce. As a thumb rule, higher the physical dimension of the workpiece, higher should be the wattage rating of the soldering iron. Some of the suggested wattage choices are given below:

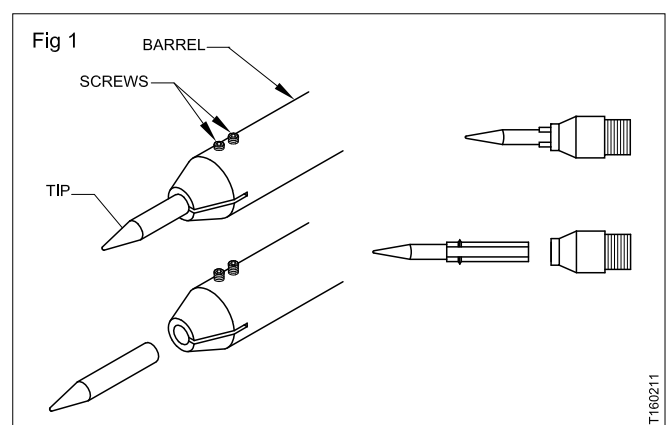
- i) For soldering less temperature sensitive components such as, resistors on lug boards, tag boards, use 25 to 60W iron. For soldering on printed circuit boards, use 10 to 25 W iron.
- ii) For soldering highly temperature sensitive components such as, diodes, transistors and integrated circuits, use 10 to 25 watts iron.

1.2 Selection of soldering iron tip

To ensure that the joint is heated to the required temperature ideally,

- the area of the tip face should be approximately equal to the area of the joint to be soldered
- the tip should be long enough to allow easy access to the joint.
- the tip should not be too long, as this may result in too low temperature at the tips working face.

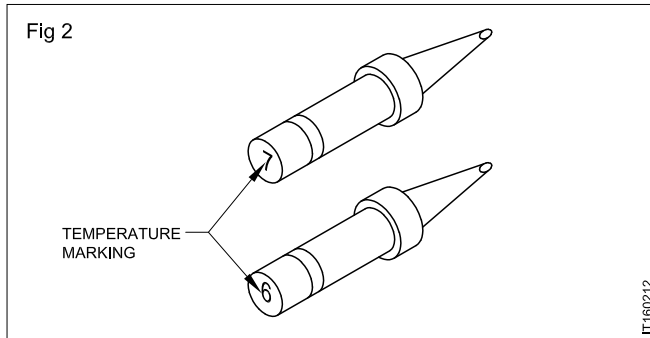
In most soldering irons, the tip can be easily removed and replaced as shown in Fig 1a and 1b.



Selection of tip temperature

Good quality soldering iron tips have numbers punched on them as shown in Fig 2. These numbers indicate the temperature to which the tip can be heated, as shown in table in the next page.

Tip No.	Temperature °C	Temperature °F
5	260	500
6	316	600
7	371	700
8	427	800



Selection of tip shape

Suggested soldering tip shapes selection table is given below;

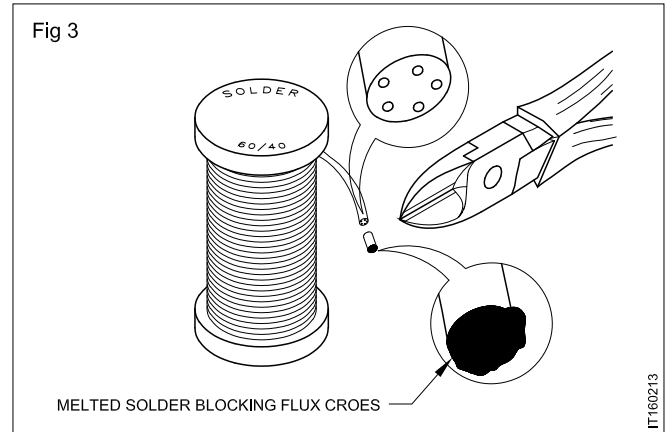
Type of soldering work	Soldering tip shape to choose
Wires, resistors and other passive components on to lug/tag boards	CHISEL TIP
All miniature electronic components except ICs on to lug boards and printed circuit boards (PCB)	BEVEL TIP
Integrated circuits (ICs) on to printed circuit boards (PCBs)	CONICAL TIP

1.3 Selection of solder and flux

As already discussed in Unit 3, there are several sizes of the cored solders whose choice depends on the size of the joints to be soldered. Also the tin and lead percentage of the solder should be checked before using the solder. Different tin and lead combinations of solder need different temperatures for it to melt and reach the liquid state.

For electronic soldering applications, solder of tin and lead of 60/40 proportion is used. This solder proportion has a melting point of 200°C which is the required temperature for general purpose soldering irons.

While soldering to make a strong solder joint the flux should melt first, and then the solder. Therefore, while using rosin cored solder, cut off the first 5 to 10mm of the solder using a side cutter (as shown in Fig 3), so that any



earlier melted portion of the solder blocking the rosin core is removed.

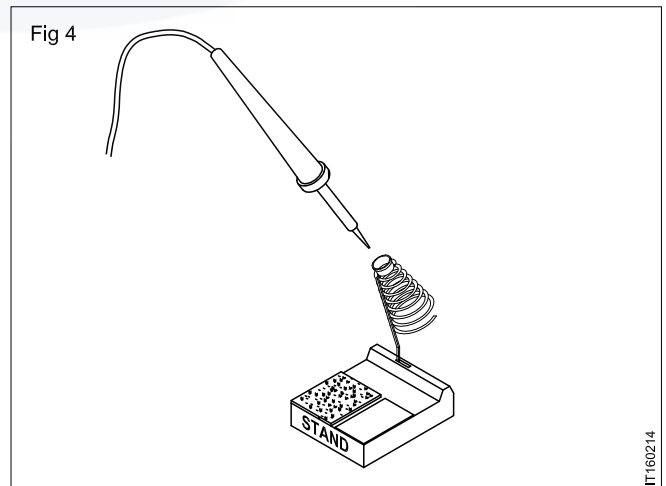
For ease of application, the flux used in addition to the cored flux in solder should be of paste form.

Flux is a chemical substance which has acidic properties. Therefore, it is advised not to touch flux by hand. Use a stick or a thin stiff brush to apply flux on workpieces. Hands should be washed after soldering work.

1.4 Soldering stand

Soldering stand plays an important role of retaining the soldering iron tip temperature around the required soldering temperature. The soldering stand should not allow the external temperature to cool the bit. At the same time the stand should not contain all the heat generated.

Soldering stands are specially designed as shown in Fig 4 to fulfill the above requirements. Such a design also prevents accidental burn injuries to the user of the soldering iron.



Another important requirement of a soldering stand is its mechanical stability. When the iron is taken out or placed in the stand frequently, the stand should not topple. An unstable stand is sure to cause burn injuries while carrying out serious soldering work.

1.5 Inspection of soldering iron

Most soldering irons are powered by AC mains voltage. This voltage level is high and can give shock if one is careless. Soldering irons will generally have lengthymains

cable. While using the iron, the mains cable gets twisted and will have to bear physical strain. Because of this strain, the insulation of cable may get cut. This may lead to live wires protruding out. The live wires give severe electrical shocks if touched by the user.

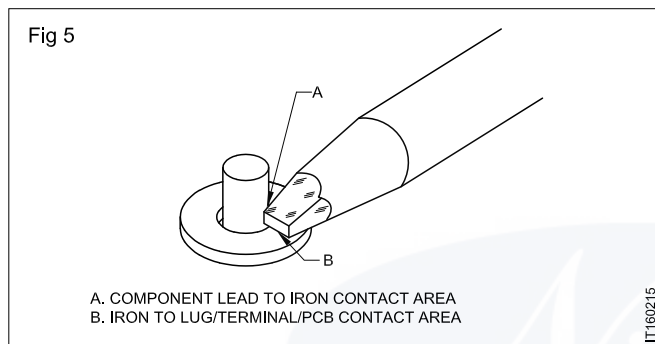
1.6 Preparation of soldering iron for soldering

Refer to Chart 1 of Ex.3.01 for the steps in preparing a soldering iron before starting the soldering work.

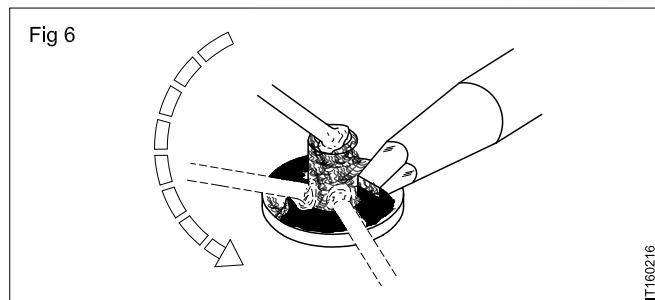
2 HEATING THE JOINT AND ADDING SOLDER

Tips for heating and applying solder to a joint to be soldered are given below:

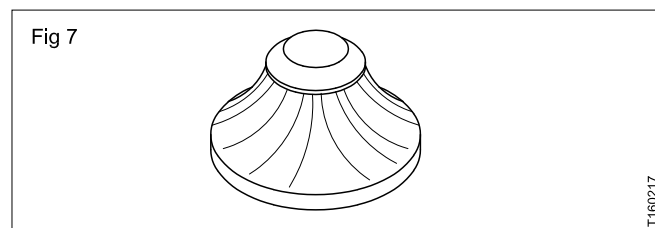
- Do not apply additional flux required for a joint in one place. Apply a small amount of flux around the joint. Do not allow the flux to flow outside the area to be soldered.
- Place the iron tip at the connection such that the tip gets maximum contact with parts to be joined as shown in Fig 5.



- Slowly feed the solder into the joint starting close to the soldering tip and moving towards the edge of the joint as shown in Fig 6.



- Continue applying the solder to the joint until complete wetting of the joint has been achieved and the joint has a concave fillet as shown in Fig 7.



- After enough solder has been applied and solder removed, keep the soldering iron tip on the joint for a moment to ensure that all the flux on the joint has reached the soldering temperature. This will allow majority of the acids within the joint to break down, which otherwise will corrode the joint after a period of time.

Generally the time taken to make a good soldered joint is between 3 to 7 seconds from applying the soldering iron to removing the soldering iron.

3 COOLING THE JOINT

Tips for cooling a solder joint are given below:

- Allow the joint to cool without assistance. Do not blow air from your mouth or from any other source to cool the joint. Forced cooling, cools the joint much earlier than it has to, resulting in a dry or brittle solder joint which will lead to mechanical and electrical defects of the joint.
- Do not move any part of the joint while it is cooling. This disturbs the chemical bonding taking place. Movement of the joint while it is cooling results in a dry joint.

4 CLEANING THE JOINT

When a solder joint is made, the amount of flux applied should be just sufficient to make a good joint. But, quite often, there will be a brown waxy substance left on the joint. This is nothing but the flux residue. In its original state this residue is corrosive. Hence, the flux residue or excess flux must be removed from the joint before soldering can be considered as complete.

If the flux residue and excess flux are not properly removed, their corrosive nature of the flux will gradually destroy the component leads and the circuit board. The flux residue is also *tacky* and, if not removed, will collect dust and debris often leading to circuit failure.

Removal of flux residue requires the use of solvents. The type of solvent depends on the flux used.

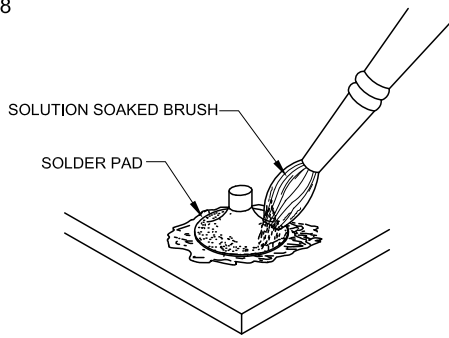
IsoPropyl Alcohol (IPA) is one of the solvents used for removing residual flux. It is available either undiluted or pre-mixed with water and can be obtained in pump sprays, aerosols, cans and drums depending on the quantity and style of use.

Cleaning using water/IPA solution

Determine the right method of application. (spray or liquid). Apply the solvent to the soldered joint. Use a clean acid brush, or some other type of stiff brush, to gently scrub the joint as shown in Fig 8, to help dissolve the residue, taking care to avoid splashing the mixture.

When the residue has been dissolved, dry the joint with a lint-free cloth to remove as much of the dissolved residue as possible.

Fig 8



IT160218

Don't 's While Soldering

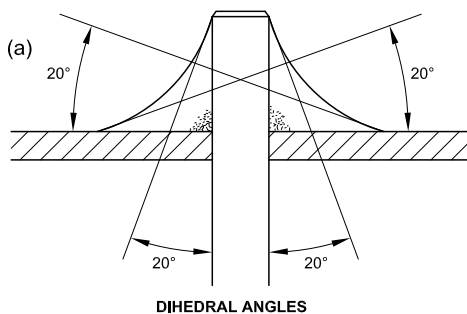
- Do not use a poorly tinned soldering tip.
- Do not cool the tip of the iron by wiping it excessively on a damp sponge.
- Do not allow the solder to be carried to the joint on the tip of the soldering iron.
- Do not attempt to speed up the cooling of the joint by blowing on it.
- Do not move the soldered joint until the solder has cooled to solid state.
- Do not try and improve a bad solder joint by reheating. All the original solder must be removed and the joint preparation and soldering should be redone.

INSPECTION OF SOLDERED JOINTS

Soon after making a solder joint, as a quick check, the following features of the solder joint should be checked:

- 1 All dihedral angles must be less than 20° as shown in Fig 9a.
- 2 All surfaces must be bright and shiny as shown in Fig 9b.
- 3 All surfaces must be smooth and symmetrical as shown in Fig 9b.

Fig 9



(b)



A GOOD SOLDER JOINT

IT160219

- 2 A closer, more detailed inspection after the joint is cleaned of residual flux.

1 INITIAL INSPECTION - Before cleaning the joint

The solder joint should be inspected only when the joint is cooled. Each joint should be examined for the following:

Shape - The shape of a correctly soldered joint will allow the profile of the underlying material to be seen. The joint will have gentle curves along its surface as shown in Fig 9b and Fig 10a. The solder will blend into a concave shape without any noticeable step at the edge of the solder joint. There will also be no sign of peaks on the surface.

Surface colour - The surface of a correctly soldered joint will be glossy, silvery and uniform in colour. There may be signs of brown waxy material (solidified flux) on the surface, but the surface underneath this will appear glossy. This excess flux must be removed using IPA or IPA solution.

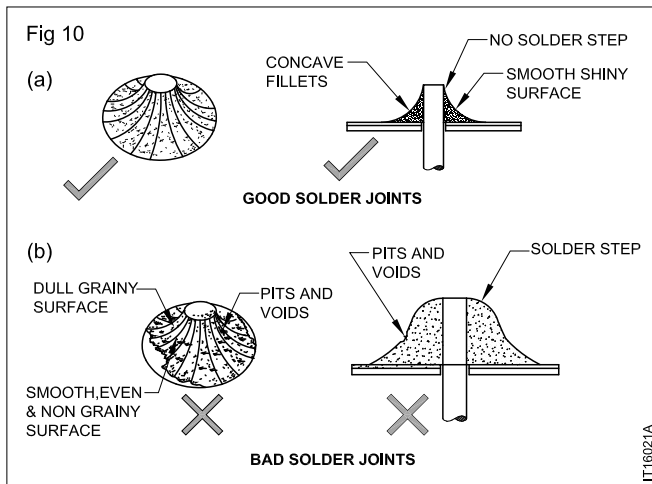
Surface texture - The surface texture of a correctly soldered joint will be smooth, even and non-grainy. There shall be no signs of pitting on the surface of the solder as shown in Fig 10.

If a solder joint does not have above features, then the joint is said to be a bad joint. Such joints will not have the required electrical and mechanical characteristics.

STAGES OF INSPECTING SOLDERED JOINTS

A well formed soldered joint will have a definite appearance. This can be assessed in two stages.

- 1 The initial appearance immediately after the joint has solidified and before cleaning the joint.

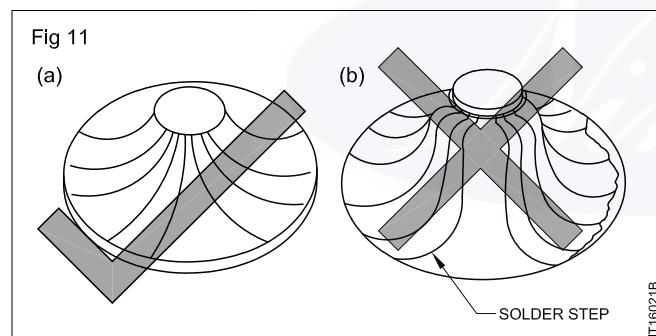


2 CLOSER INSPECTION - After cleaning the joint

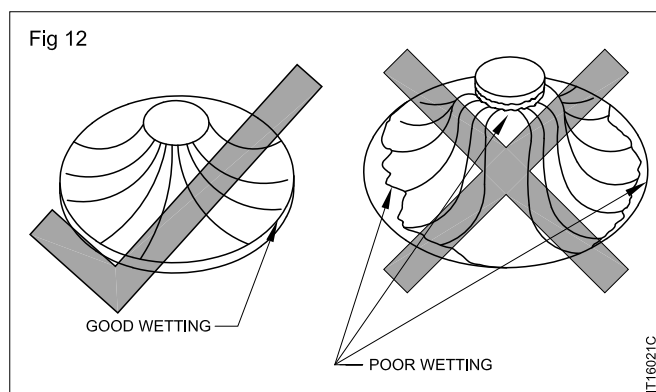
Although the initial inspection will show most of the common defects with solder joints, there are some defects which can only be identified after the flux residue has been removed. This type of inspection is best carried out using a x10 magnifying glass under a bright white light.

The features to be examined closely are listed below.

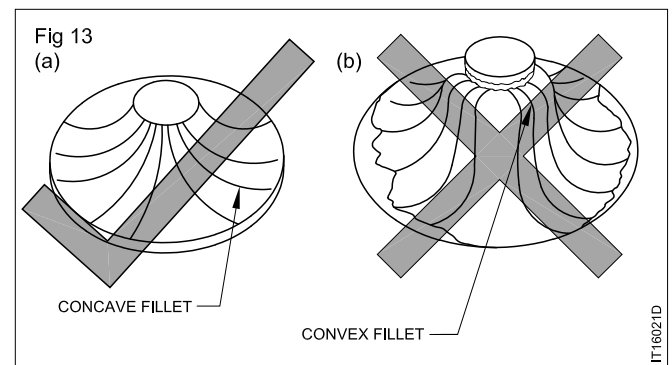
- There must be no sign of stepping at the edge of the solder flow as shown in Fig 11a. Solder step shown in Fig 11b indicates bad wetting and bad bonding between the solder and the metals.



- As shown in Fig 12a, there must be complete wetting of all protruding parts of the component lead or at least 95% of the working surface of the connection. This indicates good wetting and good bonding between the solder and the joined parts.



- As shown in Fig 13a, there must be concave fillets and not convex fillets as shown in Fig 13b. Concave fillet indicate that, correct quantity of solder has been used.



- The joint must have a smooth, shiny, light gray surface. This indicates that the solder has not been contaminated and that the joint has not been overheated. It also indicates that there has not been any movement of the joint parts while the joint cooled.
- There must be no visible damage, such as cuts, scratches, dents or fractures to the solder or component lead. This indicates that the joint is mechanically sound.

COMMON SOLDERING DEFECTS

Defects in solder joints can be grouped as follows:

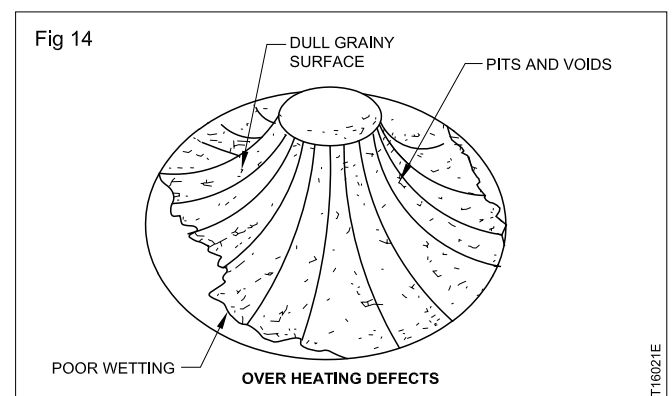
- 1 Temperature defects
- 2 Wetting defects
- 3 Solder quantity defects
- 4 Mechanical defects.

1 TEMPERATURE DEFECTS

Temperature defects are caused by excessive or insufficient heating of the joint during the soldering process.

1.1 Defects due to excessive heating

The solder on an overheated joint will have grainy texture, dull gray colour, pitting and as shown in Fig 14.



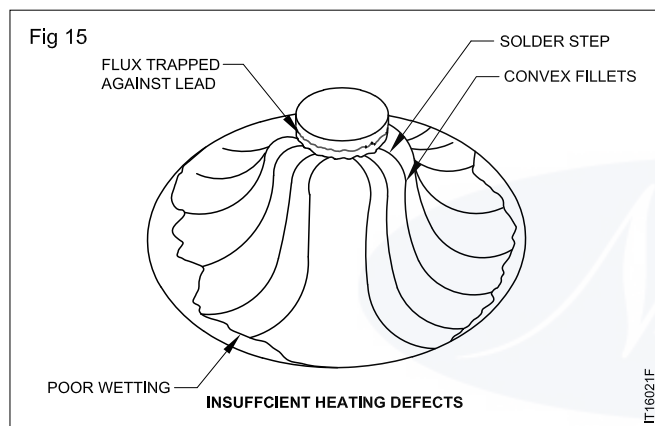
Too much heat can lead to one or more of the following defects:

- Overheating of the rosin flux, causing it to char and lose its ability to remove oxides. Lumps of charred flux gets trapped in the solder forming *pits and voids*.
- Overheating of the solder, causing excessive alloying between the copper parts of the connection and the tin content of the solder. This in turn causes localized depletion of the tin and results in a *brittle joint*.
- Excessive oxidation of the solder. The oxidised solder forms a poor bond with the other parts of the joint .

1.2 Defects due to insufficient heating

Too little heat can lead to one or more of the following defects:

A defect known as a **cold joint**. A cold joint occurs when the flux is unable to remove the tarnish from the joint. At low soldering temperature the flux is only partly activated. Therefore, it is less effective in removing tarnish. A insufficiently heated solder joint is shown in Fig 15.



Insufficiently heated solder joint results in,

- poor wetting of the joint
- a convex solder fillet
- solder steps at the edges

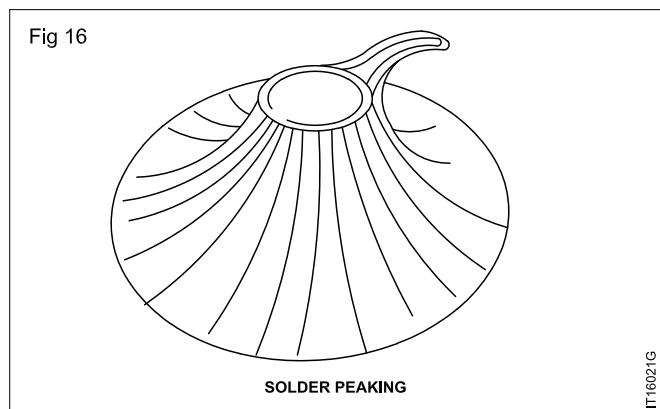
The effect of a cold joint are,

- high electrical resistance
- low mechanical strength
- dry solder joint

A defect known as a **dry joint** occurs when the solder is too viscous to push the flux away from the component lead. A layer of the flux becomes trapped around the lead as shown in Fig 15. This layer of flux causes a weak bond and hence a poor electrical connection.

Causes of incorrect soldering temperatures

A common indicator of incorrect soldering temperature is the formation of small peaks at the tip of the joint as the iron is pulled away as shown in Fig 16. These peaks will be formed at either too high or too low soldering temperatures.



Incorrect soldering temperature may be due to,

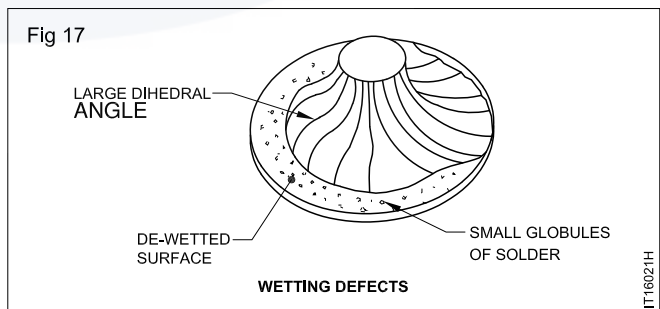
- incorrect soldering iron wattage selection
- incorrect soldering tip selection
- insufficient heating of the soldering iron
- poor soldering technique, leading to poor heat transfer from the tip of the iron to the joint.

2 WETTING DEFECTS

The degree of wetting of a joint depends greatly on the cleanliness of the parts to be joined.

Wetting defects are easily recognised by the inability of the solder to cover all the metal surfaces at the joint and the formation of a large dihedral angle as shown in Fig 17.

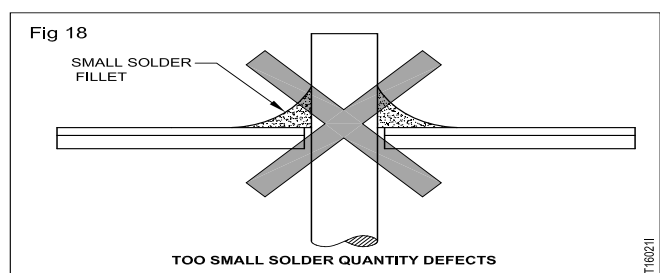
On severely tarnished surfaces, the solder will not flow over the tarnished surface (non-wetting) or the solder initially flows over the tarnished surface, but then retreats leaving small globules of solder on the surface (de-wetting) as shown in Fig 17.



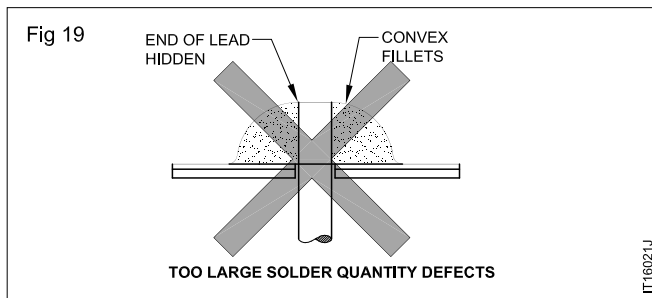
3 SOLDER QUANTITY DEFECTS

Solder quantity defects are due to either too little or too much solder applied to a connection.

Too little solder results in a small size of the solder fillet as shown in Fig 18. The small solder fillets makes a joint weak.



Too much solder results in convex solder fillets as shown in Fig 19. This defect is recognised by the large size and the convex shape of the fillets.



4 MECHANICAL DEFECTS

Mechanical defects of a solder joint are caused by;

- the movement of parts of a joint while the solder is cooling.

Movement during cooling causes severe dislocation of the crystalline structure of the solder. This results in a weak joint which may fracture later and cause high electrical resistance or an intermittent fault when the circuit is in use. A joint moved before cooling will have a frosted

appearance with fractures.

- stresses imposed on the joint as it cools.

Stresses on the joint are usually a result of inadequate stress relief bends in the leads of components. A stressed joint usually breaks due to the expansion and contraction of components while in use, due to temperature variations.

Blowing on the soldered joint to attempt to speed up cooling will introduce several additional cooling stresses.

Connections that have been disturbed as they cool usually have a frosted appearance.

INSPECTION SHEETS

Inspection sheets are used to ensure that inspection of the joint is made against a known or locally developed standard.

Based on the understanding of the information contained in the preceding paragraphs, it is advised to make a personal soldering inspection sheet and use this when checking the solder work.

Never forget the rule, If the quality or reliability of the joint is in doubt, de-solder the joint and re-solder it fresh.

De-soldering and De-soldering tools

Objectives: At the end of this lesson you shall be able to

- list the two stages in disconnecting a component from a soldered circuit
- list different tools used for de-soldering
- state the necessity for periodic cleaning of the plunger de-soldering tool
- state the common application of wicking braid.

DESOLDERING

Many a time it may be necessary to disconnect/remove components and wires from a soldered or wired circuit due to the following reasons;

- Component failure(open,short etc).
- Incorrect component installation(polarity,position etc).
- Faulty or defective solder connections(dry solder etc).
- Circuit modifications(replacing,removing components etc).

Disconnecting a component or wire from any soldered circuit involves two separate actions. These are:

- 1 DE-SOLDERING THE CONNECTION - this action involves removal of the solder from a joint
- 2 REMOVAL OF THE COMPONENT -this action involves removing the component lead from the joint.

De-soldering the connection

De-soldering is a process of heating a soldered joint, to melt the existing solder and removing the molten solder from the joint.

De-soldering makes it easy to separate or pull-out the components,wires from the joint without unnecessary damage to the components and wires.

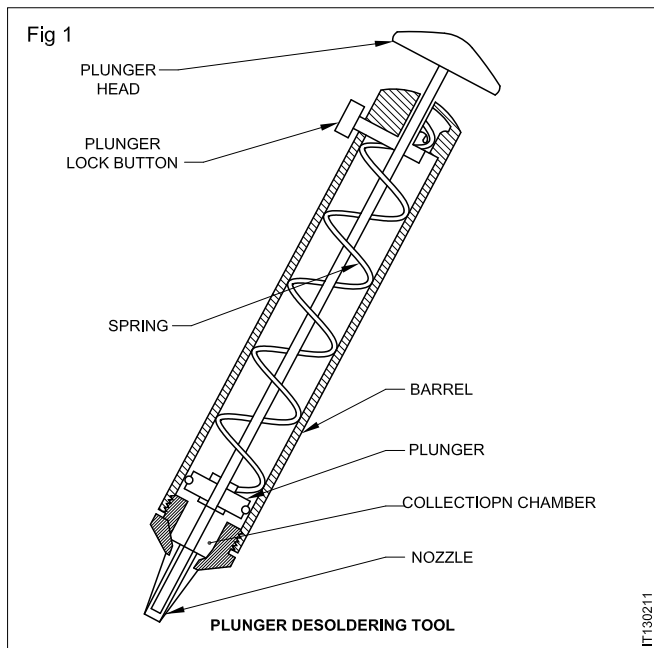
The heat required to melt the solder is supplied by a soldering iron. But removal of the molten solder from the joint requires the use of one of the following;

- Plunger de-soldering tool or desoldering pump
- Wicking braid

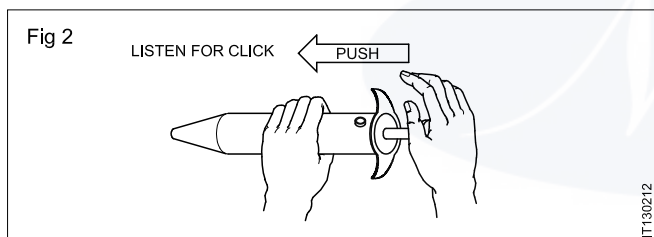
But, in many cases, desoldering is done using a nose player and a soldering iron. First, the joint to be disconnected is heated using the soldering iron. Once the solder at the joint melts, the component lead is pulled away using a nose player. This method of desoldering can be used for heavy components with strong leads. But this method should not be used for desoldering thin lead delicate components such as transistors, integrated circuits etc., This is because, in this method there is likelihood of component getting overheated or the leads getting cut or leads getting detached from the body of the component.

PLUNGER DE-SOLDERING TOOL

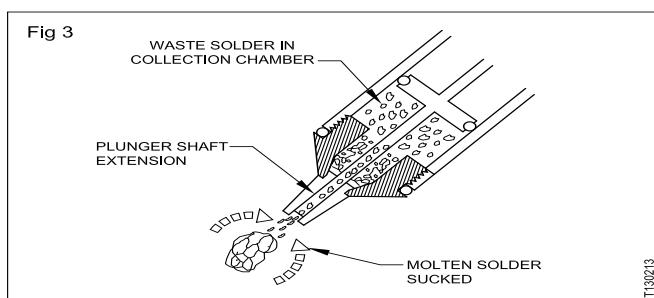
A typical plunger de-soldering tool is shown in Fig 1.



Plunger type desoldering tool is the most commonly used desoldering tool. This tool works on the principle of air suction. When the plunger head is pushed fully inside as shown in Fig 2, it gets locked inside with the help of the plunger button. This is known as cocking tool.

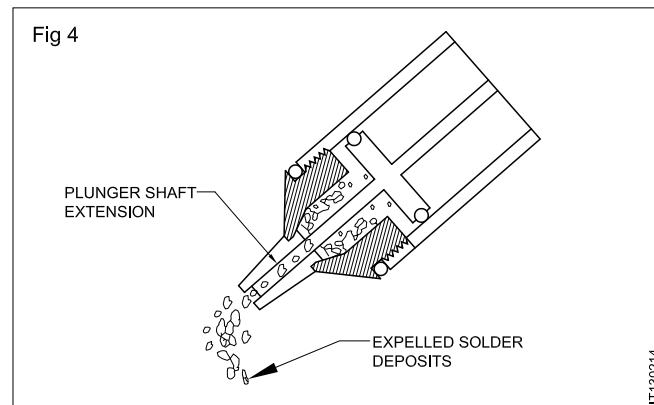


In this condition, the nozzle of the desoldering tool is kept almost touching the joint to be desoldered. If the joint is heated, the solder at the joint melts. If the plunger button of the desoldering pump is pressed, it releases the spring tension and moves the plunger up with a jerk. This causes the air to be sucked-in through the nozzle. Since the nozzle is now in contact with the molten solder, the molten solder is also sucked-in through the nozzle and gets collected in the collection chamber as shown in Fig 3.

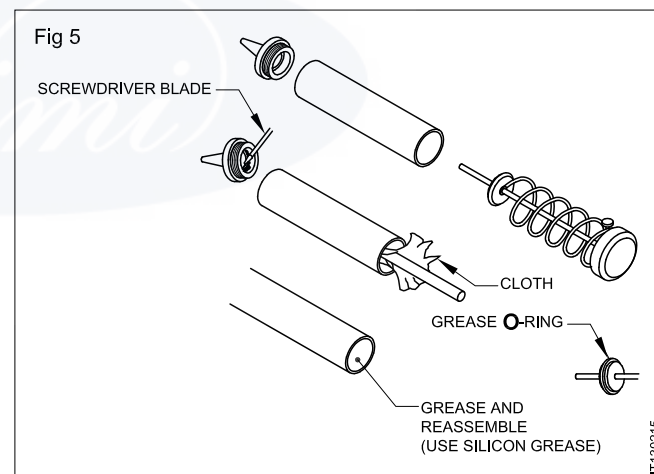


When the solder is removed using a plunger de-soldering tool, all the molten solder of a joint may not be sucked by the de-soldering tool at the first attempt, the joint must be reheated and the solder removed in two or three attempts.

After doing one suction of molten solder, while cocking the tool for second suction, face the nozzle into a dirt collector. This is because, the solder collected at the tip of the nozzle gets pushed out every time the tool is cocked as shown in Fig 4.

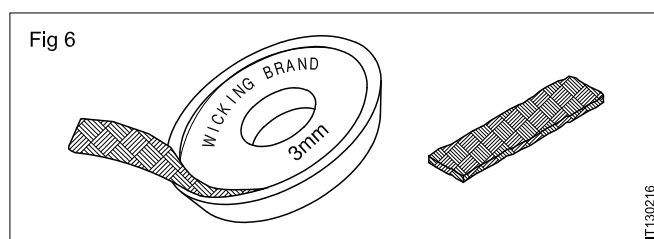


After several operations, the waste solder collected within the tool will begin to interfere with its operation. To prevent clogging of nozzle, this solder must be removed periodically and the tool must be cleaned and lubricated as shown in Fig 5.

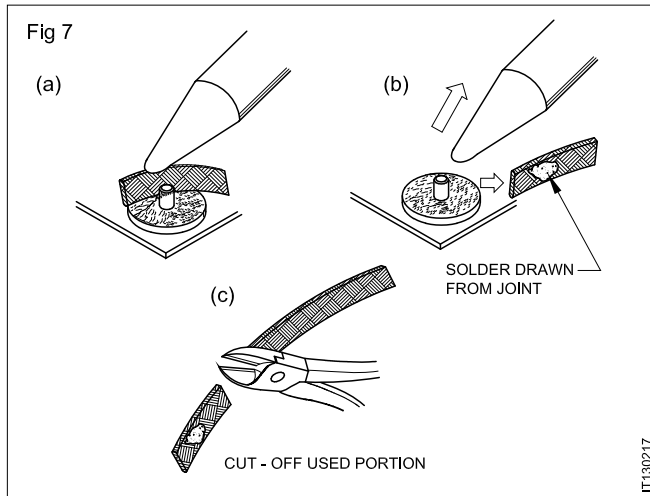


WICKING BRAID

Wicking braid as shown in Fig 6 is another simple de-soldering aid. This is made of copper and is soaked in flux. Wicking braid is nothing but a tape made of thin strands of copper knitted to form a mesh as shown in Fig 6.



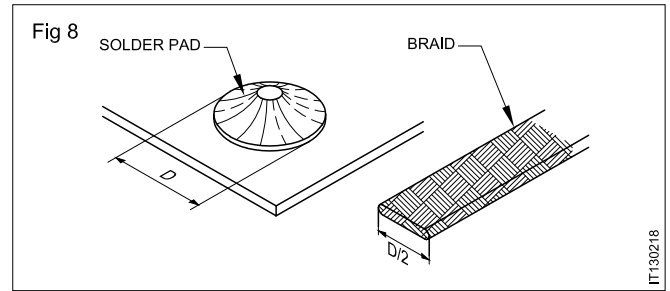
A wicking braid relies on the tendency of the hot solder to flow towards the heat source. When a soldered joint is heated via a wicking tape as shown in Fig 7a, the molten solder gets drawn into the wicking braid as shown in Fig 7b. Thus the joint is now free from solder and the component can be removed easily.



The flux content of the wicking braid varies from brand to brand. Generally, the higher the level of flux in the braid, the more efficient it will be at drawing the solder from the joint.

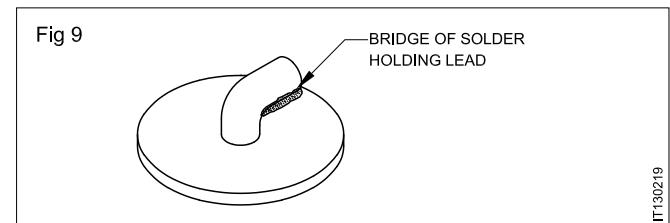
Wicking braids are available in small, hand-held rolls and is supplied in a range of sizes from 0.8 to 6 mm wide so that the correct width of wicking braid can be selected for the joint to be de-soldered as shown in Fig 8.

De-soldering using a wicking braid is commonly used for removing miniature components soldered on printed circuit boards (PCB's) and while removing components which are highly temperature sensitive.



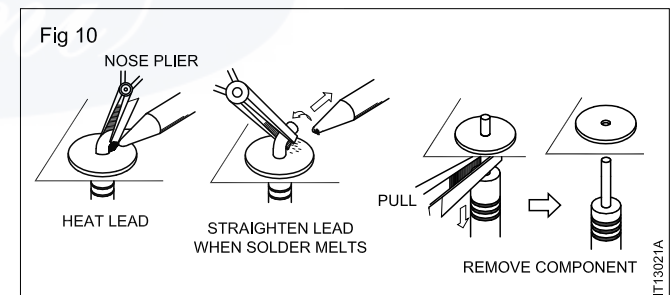
Removal of component

When solder is removed from the joint, the component can then be removed from the circuit board. If a component was soldered using clinched lead method as shown in Fig 9, it is essential to remove the bridge of solder holding the lead.



To remove the solder bridge, follow the steps shown in Fig 10

There are other special tools used for de-soldering such as De-soldering iron and multi-contact de-soldering block. For details on these tools refer reference books listed at the end of this book.



Resistors and finding its value

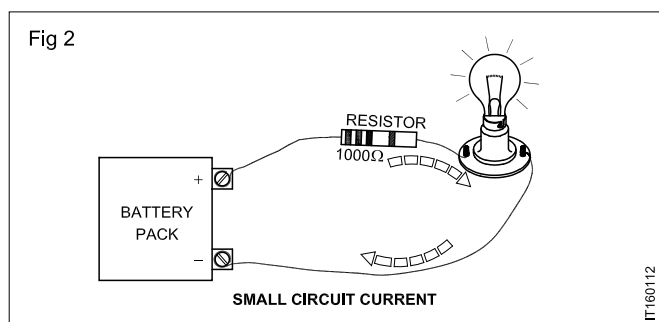
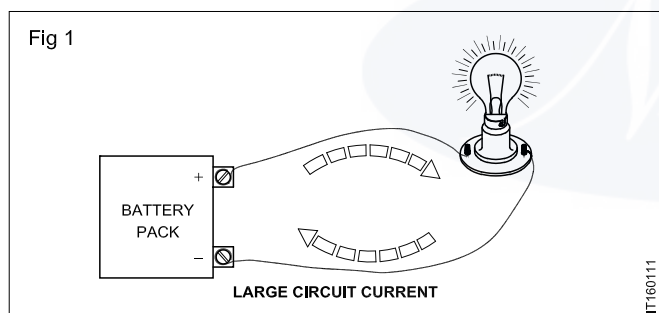
Objectives: At the end of this lesson you shall be able to

- state the function of a resistor in a circuit and unit of resistance
- name and list the classifications of resistors
- brief constructional details of important resistor types
- state the meaning of tolerance in resistor and power rating
- state the uses of an ohmmeter and list the basic components of an ohmmeter
- state the difference in the scale marking in the two main types of ohmmeters
- state the need for resistance zero setting in ohmmeters
- name the instrument to measure very accurate and high resistance values
- explain the precautions to be taken while measuring resistance.

Resistors

Resistors are electronic components, used to reduce, or limit, or resist the flow of current in any electrical or electronic circuit. Chart 1 at the end of this lesson shows different types of resistors.

Fig 1 shows a circuit in which the bulb glows brightly. Fig 2 shows the same circuit with a resistor, and the bulb glows dim. This is because, the current in the circuit is reduced by the 1000 ohms resistor. If the value of this resistor is increased, current in the circuit will be further reduced and the light will glow even dimmer.



Resistors are made of materials whose conductivity fall in-between that of conductors and insulators. This means, the materials used for making resistors have free electrons, but not as many as in conductors. Carbon is one such material used most commonly for making resistors.

When a large number of electrons are made to flow through a resistor, there is opposition to the free flow of electrons. This opposition results in generation of heat.

Unit of RESISTANCE

The property of the resistor to limit the flow of current is known as *resistance*. The value, or quantity of *resistance* is measured in units called **ohms** denoted by the symbol Ω .

Resistors are called *passive devices* because, their resistance value does not change even when the level of applied voltage or current to it is changed. Also, the resistance value remains same when the applied voltage is AC or DC.

Resistors can be made to have very small or very large resistance. Very large values of resistances can be represented as given below;

$$1000 \Omega = 1 \times 1000 \Omega = 1 \times \text{kilo} \Omega = 1 \text{ K} \Omega$$

$$10,000 \Omega = 10 \times 1000 \Omega = 10 \times \text{kilo} \Omega = 10 \text{ K} \Omega$$

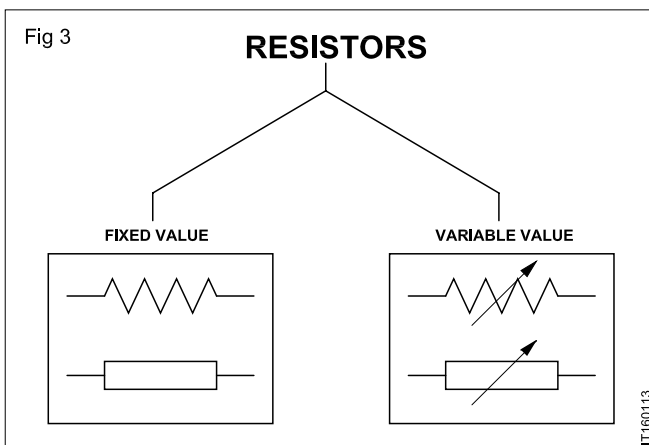
$$100,000 \Omega = 100 \times 1000 \Omega = 100 \times \text{kilo} \Omega = 100 \text{ K} \Omega$$

$$1000,000 \Omega = 1000 \times 1000 \Omega = 1000 \times \text{kilo} \Omega = 1000 \text{ K} \Omega$$

$$1000 \text{ k} \Omega = 1 \text{ M} \Omega$$

Classification of Resistors

Resistors can be classified under two main categories as shown in Fig 3.

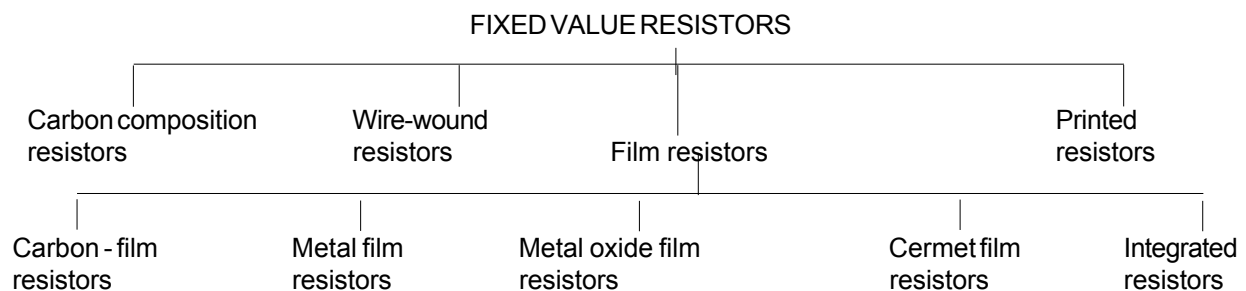


Its ohmic value is fixed. This value cannot be changed by the user. Resistors of standard fixed values are manufactured for use in majority of applications.

Fixed resistors are manufactured using different materials and by different methods. Based on the material used and

their manufacturing method/process, resistors carry different names.

Fixed value resistors can be classified based on the type of material used and the process of making as follows.

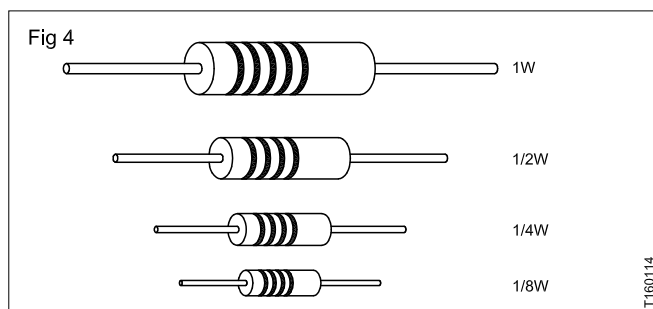


Physical appearance of some types of fixed value resistors is shown in Chart 1 at the end of this lesson.

Power rating

As already discussed, when current flows through a resistor, heat is generated. The heat generated in a resistor will be proportional to the product of applied voltage (V) across the resistor and the resultant current (I) through the resistor. This product VI is known as *power*. The unit of measurement of power is *watts*.

The physical size of a resistor should be sufficiently large to dissipate the heat generated. The higher the physical size, the higher is the heat that a resistor can dissipate. This is referred to as the power rating or wattage of resistors. Resistors are manufactured to withstand different power ratings. Fig 4 illustrates comparative physical sizes of different wattage resistors. If the product of V and I exceeds the maximum wattage a resistor can dissipate, the resistor gets charred and loses all its property. For instance, if the applied voltage across a 1 watt resistor is 10 volts resulting in 0.5 Amps of current through the resistor, the power dissipated (VI) by the resistor will be 5 watts. But, the maximum power that can be dissipated by the 1W resistor is much less. Therefore, the resistor will get overheated and gets charred due to overheat.



Hence, before using a resistor, in addition to its ohmic value, it is important to choose the correct wattage rating. If in doubt, choose a higher wattage resistor but never on the lower side. The power rating of resistors are generally printed on the body of the resistor.

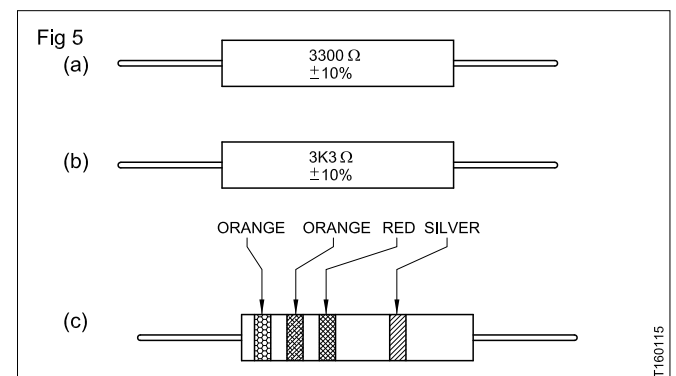
Resistor values - coding schemes

For using resistors in circuits, depending upon the type of circuit in which it is to be used, a particular type, value and wattage of resistor is to be chosen. Hence before using a resistor in any circuit, it is absolutely necessary to identify the resistor's type, value and power rating.

Selection of a particular type of resistor is possible based on its physical appearance. Chart 1 at the end of this lesson illustrates the physical appearance of most commonly used fixed value resistors. The resistance value of a resistor will generally be printed on the body of the resistor either directly in ohms as shown in Fig 5A or using a typographic code as shown in Fig 5B or using a colour code as shown in Fig 5C.

Colour band coding of resistors

Colour band coding as shown in Fig 5C is most commonly used for carbon composition resistors. This is because the physical size of carbon composition resistor is generally small, and hence, printing resistance values directly on the resistor body is difficult.



Tolerance

In bulk production/ manufacturing of resistors, it is difficult and expensive to manufacture resistors of particular exact values. Hence the manufacturer indicates a possible variation from the standard value for which it is manufactured.

This variation will be specified in percentage tolerance. Tolerance is the range(max -to- min) within which the resistance value of the resistor will exist.

Appendix D gives a list of commercially available standard resistor values.

Refer to the Appendix D for methods to read the value of resistors and their tolerance for resistors using 3 band, 4 band and 5 band colour coding schemes.

Typo graphical coding of resistors

In the typographical coding scheme of indicating resistance values, the ohmic value of the resistor is printed on the body of the resistor using an alpha-numeric coding scheme. A few examples of the generally used typographical codes and their meanings with examples are given in Appendix D.

Some resistance manufacturers use a coding scheme of their own. In such cases it will be necessary to refer to the manufacturer's guide.

Applications

Carbon composition, fixed value resistors are the most widely used resistors in general purpose electronic circuits such as radio, tape recorder, television etc. More than 50% of the resistors used in electronic industry are carbon resistors.

Brief constructional details of a few important types of fixed value resistors is given in Chart-2 at the end of this lesson.

Measuring ohmic value of resistors

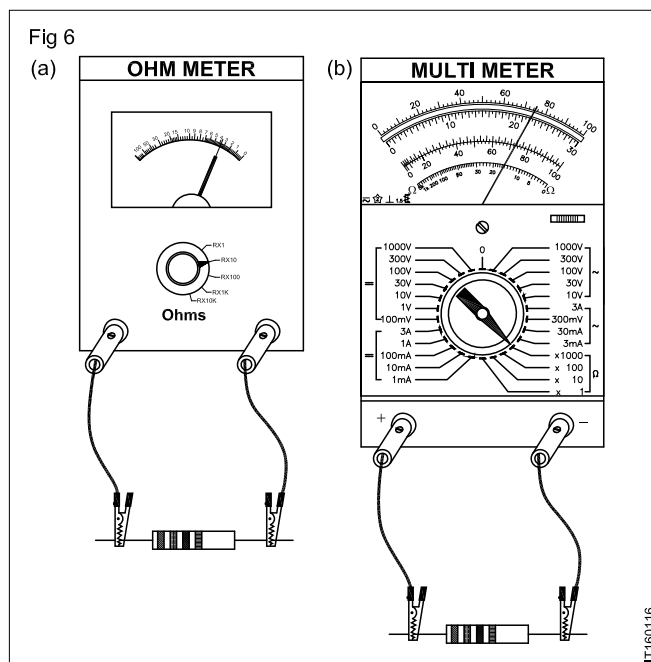
It is not possible to read the *exact ohmic value* of a resistor from colour/other coding schemes due to manufacturing tolerance built into the resistors. To find the exact ohmic value of resistors *ohmmeters* are used. When a resistor is placed between the test prods of an ohmmeter as shown in Fig 6A, the meter shows nearest to the exact resistance of the resistor directly on the graduated meter scale. Multimeters are also used to measure the value of resistors as shown in Fig 6B.

When a multimeter is used for resistance measurement, the resistance range switch on the meter should be put to the most suitable resistance range, depending upon the value of resistance being measured.

Appendix D suggest the meter ranges for measuring different resistor values accurately.

Wire-wound Resistors

Resistors, in addition to having a required ohmic value, should also be capable of dissipating the heat produced. Carbon by its nature has a limitation in the maximum heat it can dissipate. Carbon resistors become too hot when high current flows through them. This increased heat in carbon resistors changes the ohmic value of the resistors. Sometimes the resistors may even burn open due to excessive heat. Hence carbon resistors are suited only in low power circuits safely up to 2 watts.



This limitation in carbon resistors can be overcome by using wires of resistive materials like Nichrome, Manganin etc., instead of carbon. Resistors made using wires of resistive materials are known as *wire-wound* resistors. These resistors can withstand high temperature, and still maintain the exact ohmic values. In addition, wire-wound resistors can also be made to have fractional ohmic values which is not possible in carbon composition resistors.

Resistor values

Wire-wound resistors are available from a fraction of an ohm to 100's of Kilo ohms, with a power ratings of 1 watt to several 100s of watts. The higher the power rating, the thicker the resistive wire used, and bigger will be the physical size of the wire-wound resistor.

Applications

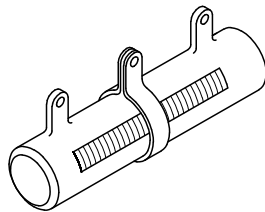
Wire-wound resistors are commonly used in electronic circuits where small values, precision values, high wattage ratings are required. A few applications are : regulated power supplies, amplifiers, motor controls, servo control circuits, TV receivers etc.

Special types of fixed value wire wound resistors

In applications where more than one fixed value wire-wound resistor is required to be used, a tapped wire wound resistor with more than one value, made in a single unit as in Fig 7 can be used.

Tapped resistors, whose tapings can be adjusted by adjusting the position of the sliding collar are also available as shown in Fig 8. This gives the flexibility of varying the resistance value between the tapings.

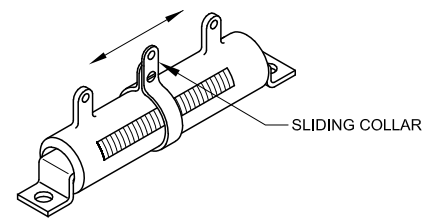
Fig 7



TAPPED WIRE WOUND RESISTOR

IT160117

Fig 8



SLIDING COLLAR

IT160118

Comparison of fixed value carbon and wire-wound resistors

Characteristics	Carbon resistors	Wire-wound resistors
Size	Small	Comparatively large
Values available	From 1 ohm to several 10's of megohms	Fraction of an ohm to several 10's kilo ohms
Tolerance	Generally 5% to 20%	As low as 1% available
Wattage	Generally up to 2 watts	As high as 100 watts
Suitability in circuit	Suitable for Highfrequency circuits	Unsuitable for high frequency circuits (Suitable up to 200 KHz)
Noise generated	Less noisy	Very noisy operation
Voltage	Low voltage rating	Wide voltage rating

OTHER TYPES OF RESISTORS

Brief constructional details of some types of resistors is given in Chart 2 at the end of this lesson. For further details on these types of resistors, refer reference books listed at the end of this book.

Resistors lead bending

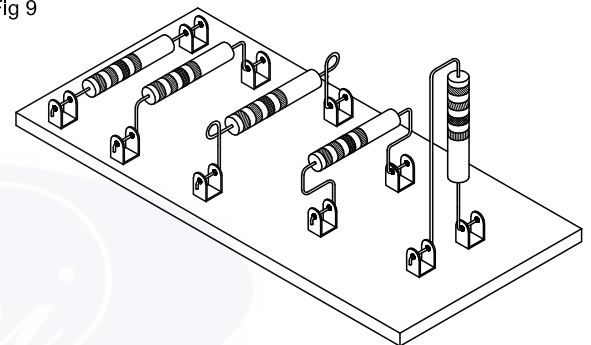
The most delicate portion of a resistor is the joint between the leads and the resistor body. The leads give electrical connection to a resistor with the external circuit. While connecting a resistor in a circuit, the leads may come across vibration, shock, rough handling etc. In spite of these, the leads must maintain perfect electrical contact with the resistor. If the lead joint is loose with the body, it may lead to intermittent electrical contact resulting in the malfunctioning of the circuit.

Resistors are the most common components in any electric and electronic circuits. The circuits may be wired on, lug boards or on printed circuit boards. When resistors are to be soldered on boards, it becomes necessary to bend the leads of the resistors to different shapes as shown in Fig 9.

While bending the resistor leads, the following two vital points should always be kept in mind. Otherwise, the strain on the lead may make the resistor's body-to-lead connection weak.

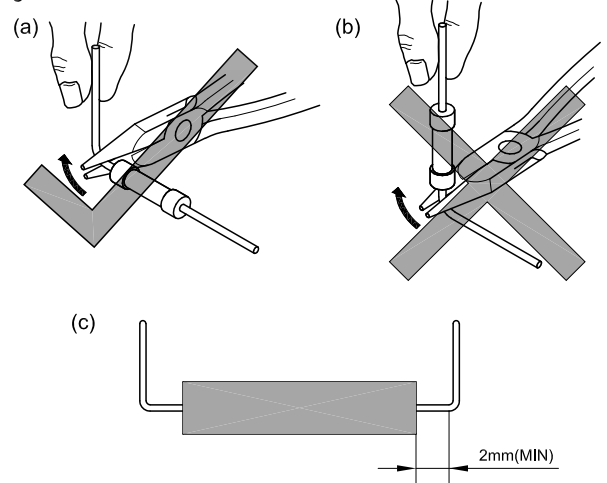
- 1 Never bend the component as shown in Fig 10b. Bend the lead as shown in Fig 10a.
- 2 Bend the lead at least 2mm away from the resistor as shown in Fig 10c.

Fig 9



IT160119

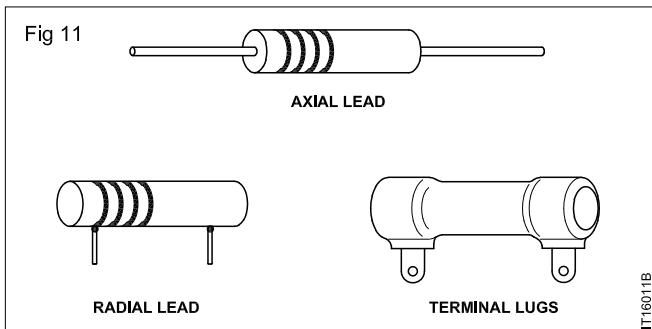
Fig 10



IT16011A

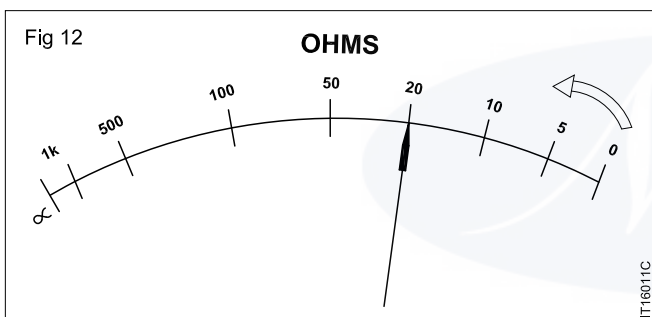
Types of resistor leads

Resistors are available with different types of lead attachment as shown in Fig 11. This makes it easy for the user to mount the resistors in different ways on lug boards, PCBs and other types of circuit boards.



OHMMETERS

A colour coded resistor indicating 22Ω with 10% tolerance can have resistance value anywhere between 24.2 ohms to 19.8 ohms. But, its exact value is still unknown. To find its exact value, a measuring instrument known as *ohmmeter* is used. These ohmmeters indicate the resistance values directly on a calibrated scale, as shown in Fig. 12.



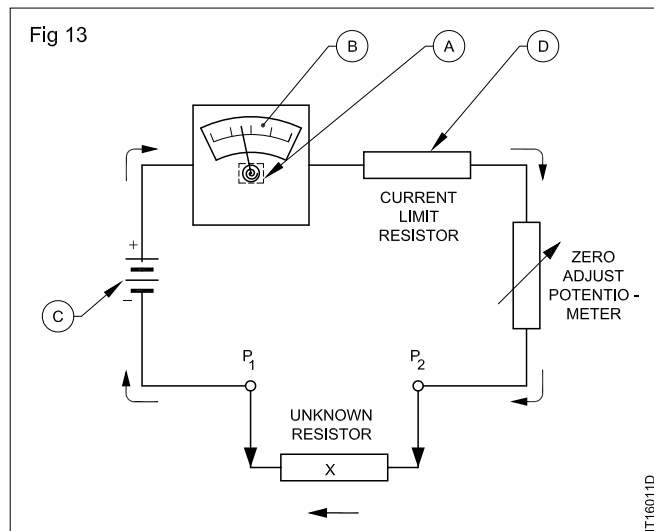
Ohmmeters can also be used to measure the resistance of any component or circuit. Since an open circuit or a short circuit means, infinite resistance and zero resistance respectively, ohmmeters can also be used to check open circuits and short circuits.

Basic components of a ohmmeter

An ohmmeter consists of a dc current meter movement(A), graduated scale(B), a low voltage dc power source(C) which will generally be a set of dry cells, a current limiting resistor(D) and a zero adjust potentiometer as shown in Fig 13.

The unknown resistance(X) whose ohmic value is to be measured is connected across the meter terminals P1 & P2. The battery makes the current to flow through the unknown resistor X. The quantity of current that flows through X depends on the ohmic value of the resistor X. This current makes the meter movement to deflect the pointer. The amount of deflection, therefore, depends on the quantity of the current flow, which in turn, is dependent upon the value of the unknown resistor X. The scale of the meter is calibrated in ohms as shown in Fig 12. Hence the pointer position on the graduated scale directly indicates the exact value of the unknown resistor X in ohms.

Fig 13



The zero adjust potentiometers as discussed in lesson 4.02, is used for *Resistance zero setting* of the ohmmeter.

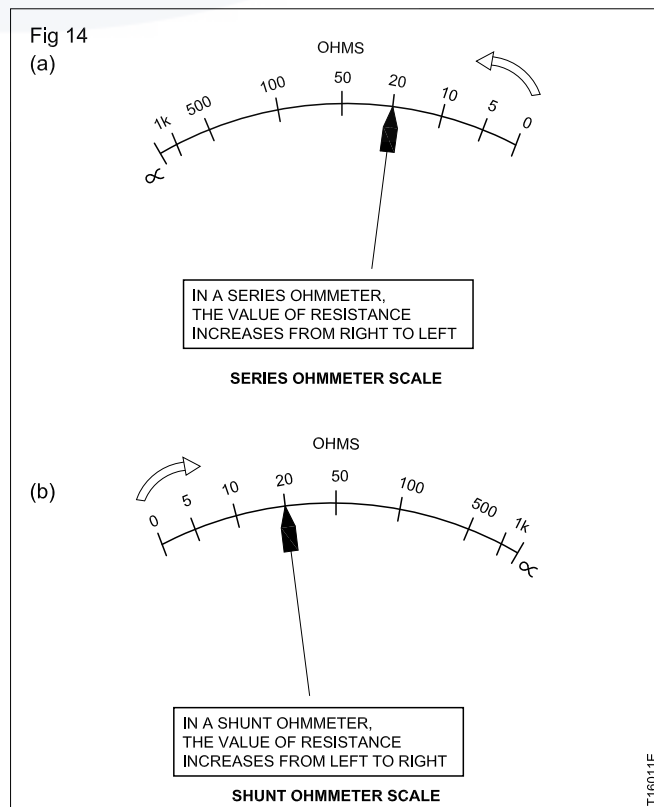
Types of ohmmeters

Depending upon the type of internal circuit used, ohmmeters are of two types;

- 1 Series ohmmeter
- 2 Shunt ohmmeter

Difference between these types of meters and their working principle is dealt in detail in further lessons. However it should be noted, that one important difference between these types of meters is the scale graduations. Scale markings of these two types of meters are shown in Fig 14a & 14b.

Fig 14



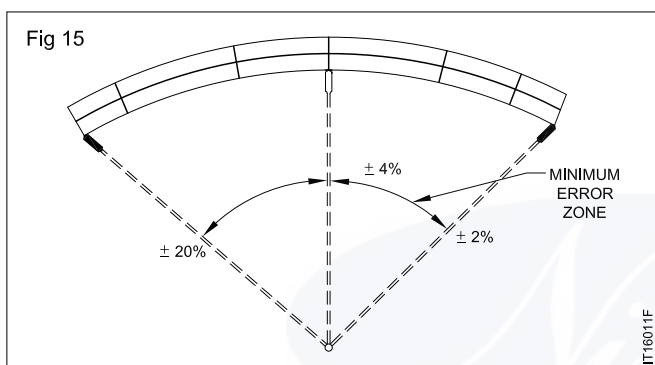
READING ERRORS IN METERS

1 Errors due to moving parts

The meter movement (label A of Fig 2) of meters depends on a magnet, spring(s), bearings and several other mechanical parts. Because these mechanical parts and pointer needle is involved in the meter movement, the meter will have measuring error. This error varies along the length of its scale as shown in Fig 15. This error will be maximum at the left edge and minimum at the right edge as shown. Hence while taking readings on a meter scale, readings will be most accurate if the reading is taken in the minimum error portion of the meter.

2 Parallax error

This error occurs in all measuring instruments employing a graduated scale and pointer. This error occurs because, while taking the reading pointed by the pointer, if the observer's eye and the pointer is not in line and perpendicular to the plane of the scale, the reading taken will be slightly different from that actually shown in Fig 15.



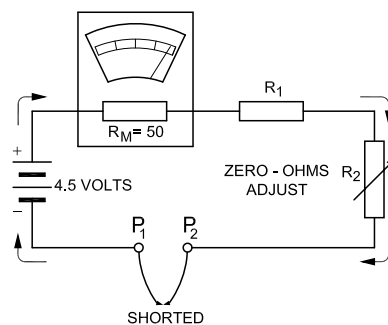
This is known as Parallax error. To minimise this error, always look into the meter perpendicular to plane of the scale while taking readings.

3 Zero setting errors in ohmmeters

In addition to the above, since ohmmeters need an internal battery, it is necessary to check the internal battery voltage before using the meter. Low battery voltage adversely affects the accuracy of resistance measured. Hence a procedure called resistance zero setting of the ohmmeter is carried out before measuring an unknown resistance value.

Resistance zero setting of the meter means, checking full scale deflection of the meter with the meter probes shorted as shown in Fig 16. The dry battery inside the meter becomes weak with time and use. As they age, their output voltage drops. As the voltage drops, the current it can supply decreases and the meter will no longer deflect full scale. Under this circumstance, the readings taken on the meter will not be accurate. To accommodate for small variations in the battery voltage, a potentiometer (R_2 of Fig 16) is connected in series with the meter circuit. By decreasing the value of this resistance, the current in the meter circuit can be increased, and hence, full scale deflection of the meter movement can be achieved. If it is not possible to get full scale deflection by adjusting R_2 , it is necessary to replace the batteries with a new set, to get accurate readings.

Fig 16

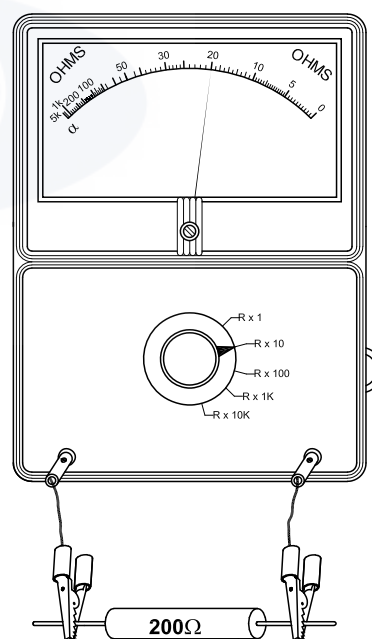


Choosing ohmmeter range

As discussed in lesson 4.02, to enable measurements of a wide range of resistance values (ohms to several megohms) multi-range ohmmeters as shown in Fig 17, with ohms multipliers are available. Depending upon the value of resistance to be measured, the most suitable multiplier ranges can be chosen such that the meter readings is taken between the mid-portion and right edge of the graduated scale where the meter error will be less.

For example as shown in Fig 17, to measure a resistance value of say 200 ohms, although this value can be measured in $\times 1 \Omega$ range, it is better to put the multiplier switch to $\times 10 \Omega$ range as the pointer position in this range will be in minimum error zone of the scale as shown in Figure 17.

Fig 17

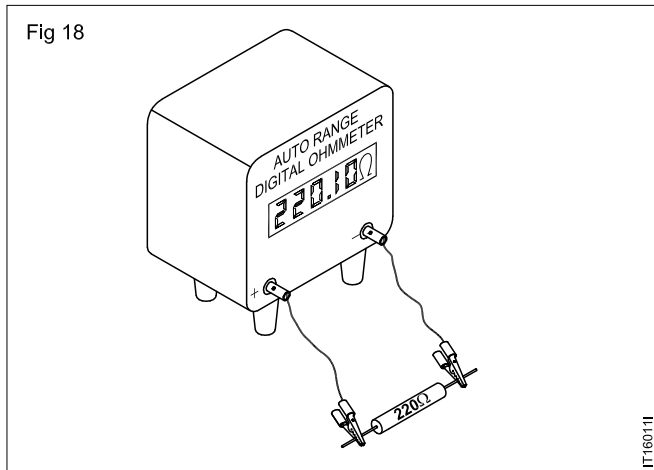


Meter errors, discussed above, limit the accuracy of analog ohmmeters. This is especially true while measuring resistance of very low (less than 1 ohm) and very high (more than a few hundred megohms) values.

Digital ohmmeters

Digital meters work on the principle of *digital electronics* (discussed in further lessons). These meters will not have any movable parts such as the meter movement, pointer etc., as in analog meters.

A typical digital ohmmeter is shown in Fig 18. Digital meters are more accurate than analog ohmmeters because of the following reasons;



- No mechanical movement is involved.
- No parallax error (meter reading error) occurs as ohmic values are shown directly in numerals.
- Can show values with two or three decimal place accuracy directly.

Digital meters are more expensive than analog meters.

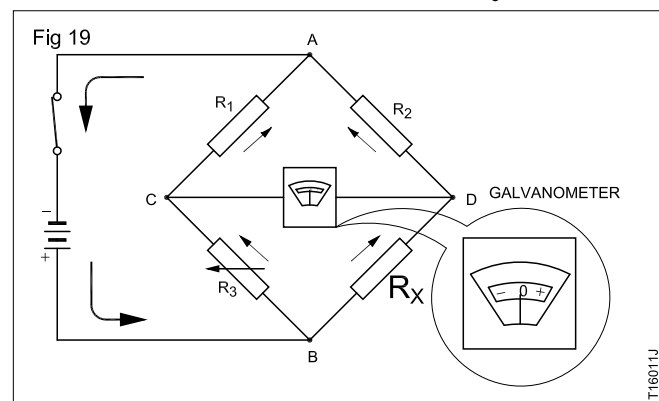
Multimeters as ohmmeter

As already discussed in lesson 4.02, multi-meters are used to measure several electrical parameters like voltage, current and resistance. These meters are more popular because of their integrated measurement capabilities.

In these meters, for measuring resistance, the control switch should be put to the suitable resistance range depending upon the value of resistance being measured.

WHEATSTONE BRIDGE

Wheatstone bridge is another commonly used instrument to measure accurate resistance values. A Wheatstone bridge, shown in Fig 19, consists of four resistors connected in a diamond shaped array, a center zero galvanometers and a dc power source. One of the resistors in the diamond shaped array will be the unknown resistance (R_x), whose value is to be measured. While making measurement the value of R_3 in Fig 19 is adjusted till the galvanometer shows zero reading. With this adjustment, the value of the unknown resistor will be equal to that of R_3 .



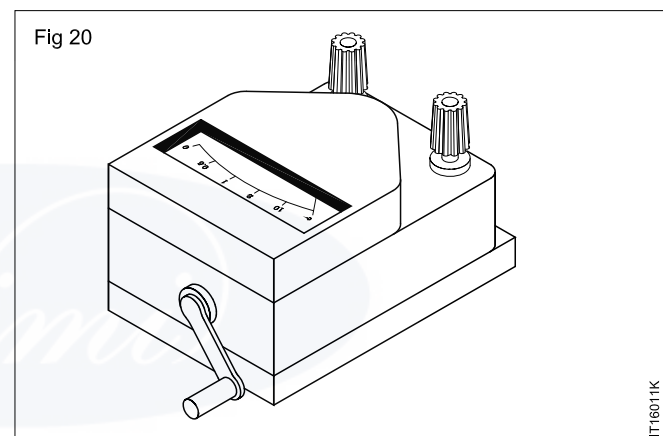
Measurement of resistance using a Wheatstone bridge is more time consuming than using a direct measuring ohmmeter but, the accuracy of the measured values will be very high.

Wheatstone bridges are not commonly used to measure resistance smaller than 1 ohm or greater than 100,000 ohms.

MEGGERs

Bridges and simple meters, both analog and digital, cannot be used to measure very high resistance values of the order of several thousands of megohms. To measure these very high resistance values such as, insulation resistance of, cables, motor winding, transformer winding etc., an instrument known as *Megohm meter* or more commonly known as *Megger* shown in Fig 20 is preferred. This is because, meggers are specially designed to measure resistance values ranging from several hundred megohms to several thousand megohms.

Meggers are most frequently used by electricians to check insulation resistance.



Precautions to be taken while measuring resistance

- If the resistance to be measured is not approximately known, set the ohmmeter to the maximum resistance range. Then, go down the ranges in steps till the resistance under measurement is shown in the upper half or middle of the meter scale.
- Touch or clip the ohmmeter terminals firmly to the resistor under measurement.

A loose contact or open may show very high resistance which may be misunderstood to be the value of the resistor under test.

- Do not touch both the conductor portion of the meter probes while making resistance measurement.

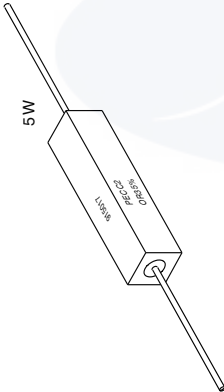
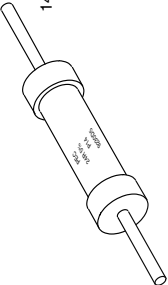
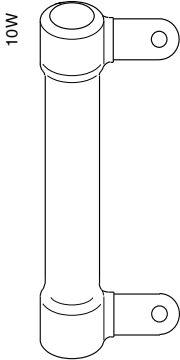
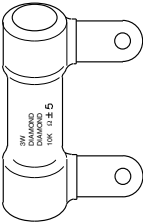
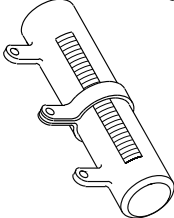
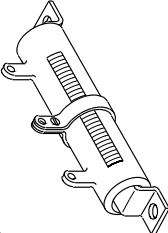
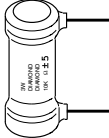
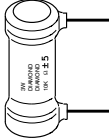
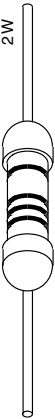
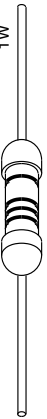
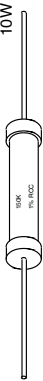
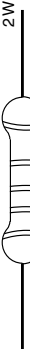
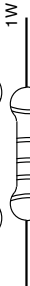
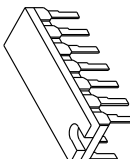
If both the terminals are touched by fingers, the meter shows the effective value of body resistance in parallel with the resistance measured and not the component resistance

- 4) Switch OFF or disconnect power supply/battery of the circuit to measure the resistance connected in the circuit.
- 5) If the resistance to be measured is connected in a circuit, de-solder or disconnect one end of the resistor before measuring its resistance.
- 6) Discharge all capacitors, particularly electrolytic capacitors, in the circuit before making resistance measurement.

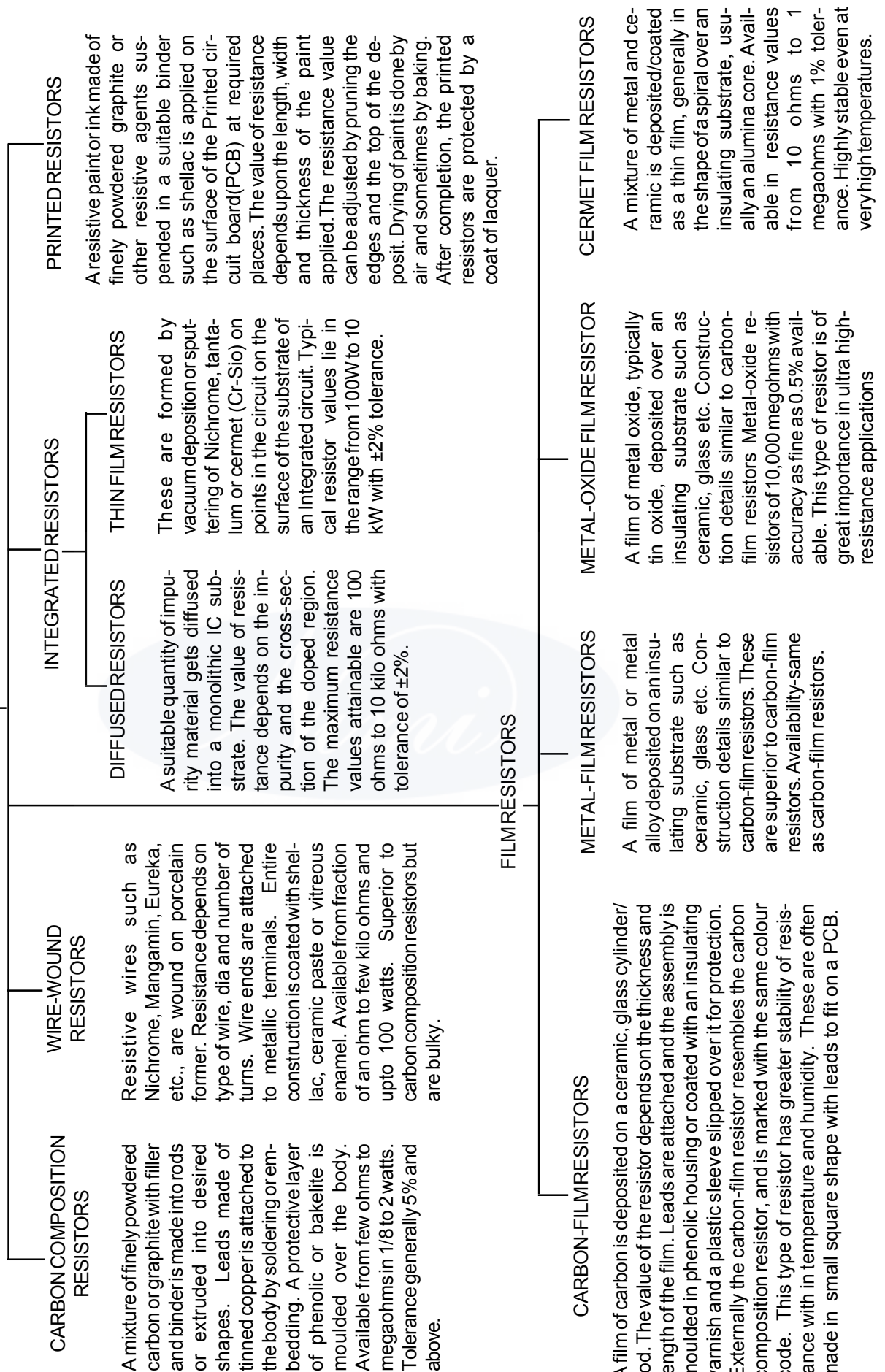
If the circuit is energised while measuring resistance, its voltage could cause a damaging current to flow through the meter.



Chart 1 for Lesson 1.2.07

FIXED VALUE RESISTORS			WIRE WOUND TYPES		SPECIAL TYPES
CERAMIC TYPES		RADIAL LEADS			
CARBON TYPES CARBON COMPOSITION	10W  5W  1W 		14W 		
	RADIAL LEADS		10W  2W     2W  1W 		
METAL FILM		RADIAL LEADS		PRECISION RESISTOR	
2W  1W  1/2W  1/4W  10W  1W 		METAL FILM RESISTOR		METAL FILM RESISTOR	
METAL OXIDE		RADIAL LEADS		LOW OHM METAL FILM	
2W  1W  1/2W 		VERTICAL MOUNT		INTEGRATED RESISTOR	
					

CONSTRUCTIONAL DETAILS OF FIXED VALUE RESISTORS



Ohm's law

Objectives: At the end of this lesson you shall be able to

- state Ohm's law
- calculate current in a circuit with one resistor
- calculate the total resistance of series resistance circuits
- calculate the power dissipated across a resistor and circuit
- design a simple voltage divider circuit using resistors
- calculate the total resistance of parallel resistance circuits
- calculate the current flowing through branches of a parallel resistance circuit
- calculate the total resistance of series-parallel resistive circuits
- calculate the resistance value required to limit current in simple circuits
- calculate and connect resistors to obtain non-standard resistance values.

OHM'S LAW

The quantity of current flowing through a resistor depends on two factors:

- 1 The ohmic value of the resistor.
- 2 The voltage applied across the resistor.

If the voltage applied across a resistor is kept constant, higher the resistance of the resistor, lower will be the current flowing through it. In other words current (I) through a resistor is inversely proportional to resistance (R) value of the resistor.

On the other hand, if the applied voltage (V) across a fixed value resistor is increased, the current flowing through the resistor also increases. In other words current (I) through a resistor is directly proportional to the applied voltage (V) across the resistor.

Combining the above two relationships between resistance (R), current (I) and applied voltage (V), it can be written as,

$$I = \frac{V}{R}$$

This relationship of $I = V/R$ was found by the scientist George Simon Ohm and hence this is referred to as ohm's law.

The relationship of $I = V/R$ can be expressed mathematically in different forms as

$$I = \frac{V}{R} \text{ or } V = I \times R \text{ or } R = \frac{V}{I}$$

These formulas are used invariably while designing or testing electrical/electronic circuits.

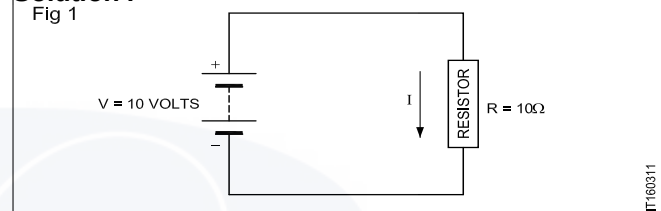
Generalising, ohm's law can be stated as follows:

Under a given constant temperature, the current flowing through a resistor is directly proportional to the voltage across the resistor and inversely proportional to the value of resistance.

This statement holds good not only for a resistor, but in common to all resistive circuits.

Example 1 : Using ohms law, find the current flowing through the resistor in Fig 1.

Solution :



Applied voltage across the resistor is : 10 volts

Resistance value of the resistor is given as 10 ohms.

Therefore current (I) through the resistor by Ohm's law is;

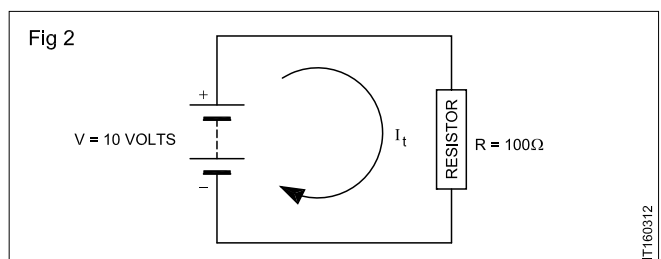
$$I = \frac{V}{R} \text{ Amps.} = \frac{10 \text{ volts}}{10 \text{ ohms}} = 1 \text{ amp.}$$

Current through the resistor is 1 ampere.

Example 2 : Find the current flowing through the circuit in Fig 2.

Solution :

Let the total circuit current be I_t .



From Ohm's law

$$I_t = \frac{\text{Total voltage applied to the circuit (V)}}{\text{Total circuit resistance (R)}}$$

In the given circuit ,

total applied voltage = 10 volts.

Total circuit resistance = resistance of the only resistor = 100Ω.

This is because the given circuit consists of one resistor and that is the only component in the circuit.

Total current through the circuit I_t is equal to

RESISTORS IN SERIES

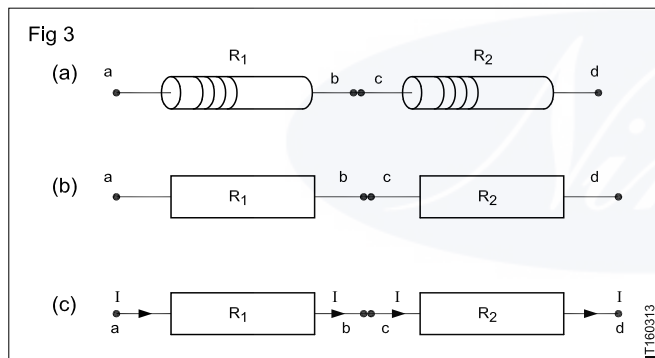
$$I_t = \frac{V}{R} = \frac{10 \text{ V}}{100 \Omega} = 0.1 \text{ amps} = 100 \text{ mA.}$$

When resistors are connected end to end as shown in Fig 3, the resistors are said to be in series with each other.

Total resistance of resistors in series

When resistors are connected in series, the total resistance of the series connection will be equal to, the sum of individual resistance values. In Fig 3, total resistance across points a-d will be equal to $R_1 + R_2$.

Example : In Fig 3, if R_1 is 1 K ohms and R_2 is 2.2K ohms. The total or effective resistance between the terminals a and d will be,



R_1 and R_2 are connected in series.

Therefore, the effective circuit resistance is

$$\begin{aligned} &= R_1 + R_2 \\ &= 1.0 \text{ K } \Omega + 2.2 \text{ K } \Omega = 3.2 \text{ K } \Omega. \end{aligned}$$

Current through a series circuit

When resistors are connected in series as shown in Fig 3, the current that flows through R_1 can only flow through R_2 . This is because

- there is no other path for any other extra current to flow through R_2
- there is no other path for the current through R_1 to escape from flowing through R_2 .

Therefore in a series circuit, the quantity of current will be the same at all the points (a,b,c,d) of the circuit as shown in Fig 3c.

The quantity of current flowing through the series path is decided by both the resistors put together or the effective resistance of the circuit.

Example : Find the total circuit current(I_t) in the circuit at Fig 4.

Solution :

Resistors R_1 & R_2 are in series. Therefore, the effective resistance of the circuit = $R_1 + R_2$

$$= 3.3\text{K ohms} + 330 \text{ ohms.}$$

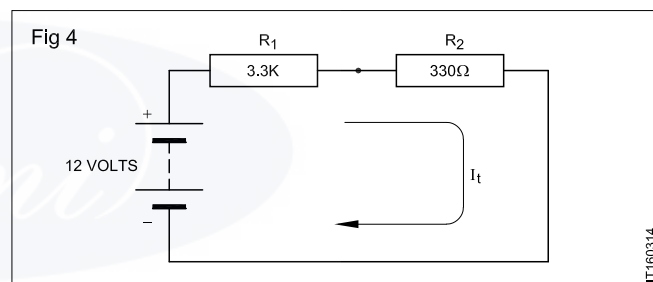
$$= 3300 + 330 = 3630 \text{ ohms.}$$

$$\begin{aligned} \text{Circuit current } I_t \\ &= \frac{V}{R} = \frac{12 \text{ V}}{3630 \Omega} = 0.0033 \text{ amps} = 3.3 \text{ mA.} \end{aligned}$$

Example : Calculate the voltage drops across R_1 and R_2 for the circuit at Fig 4.

Solution :

In the circuit (Fig 4), R_1 and R_2 are in series. Hence the current through both the resistors is the same. This current is 3.3 mA as calculated in the previous example.



From Ohm's Law

$$I = \frac{V}{R} \text{ or } V = I \times R.$$

Therefore the voltage drop across R_1

$$\begin{aligned} &= I \times R_1 \text{ volts} \\ &= 3.3 \text{ mA} \times 3.3 \text{ k}\Omega \\ &= (3.3 \times 10^{-3}) \times (3.3 \times 10^3) \\ &= 3.3 \times 3.3 = 10.89 \text{ volts.} \end{aligned}$$

Similarly the voltage drop across R_2

$$\begin{aligned} &= (3.3 \times 10^{-3}) \times 330 \text{ ohms} \\ &= 1089 \text{ milli-volts} \\ &= 1.089 \text{ volts.} \end{aligned}$$

Verification of solution

Since R_1 and R_2 are in series, the sum of the voltage drops across R_1 and R_2 must be equal to the applied battery voltage of 12V.

i.e, $10.89 + 1.089 = 11.979 \approx 12$ volts = applied battery voltage.

POWER DISSIPATION IN RESISTORS

When current flows through a resistor heat is generated. This is because, the voltage driving the current through the resistor is doing some amount of work in overcoming the opposition to the flow of electrons. It is found through experiments and analysis that, the amount of work done by the voltage is directly proportional to the ohmic value (R) of the resistor and square of the current (I^2) flowing through the resistor. This work done is dissipated in the form of heat generated by the resistor. This heat dissipating capacity is known as the power or wattage of a resistor. The unit of power is Watt.

Power dissipated by a resistor = $I^2 \times R$ Watts.

Where,

I is the current through the resistor

and R is the resistance of the resistor.

Example : If 10 mA flow through a resistor of 10 K ohms, what is the power dissipated by the resistor ?

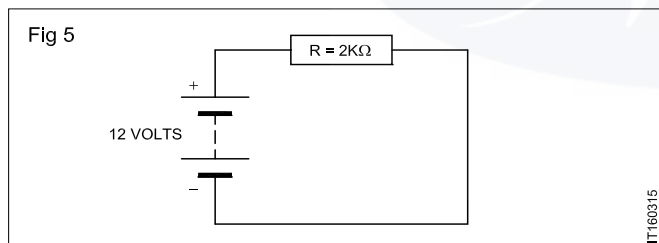
Power dissipated by the resistor = $I^2 \times R = (I \times I) \times R$

$$= (10 \times 10^{-3}) \times (10 \times 10^{-3}) \times (10 \times 10^3)$$

$$= 1000 \times 10^{-3} = 1000 \text{ milli-watts} = 1 \text{ watt.}$$

The power dissipated by the resistor is 1 watt.

Example : What is the total power dissipated by the circuit given at Fig 5.



Solution :

Method 1

Current through the circuit is $I_t = V/R$
 $= 12V / 2 K\Omega = 6 \text{ mA}$

Power dissipated by the circuit is

$$= (\text{circuit current})^2 \times \text{circuit resistance}$$

$$= (36 \times 10^{-6}) \times (2 \times 10^3)$$

$$= 72 \times 10^{-3} \text{ watts}$$

$$= 72 \text{ milli-watts} = 0.072 \text{ watts.}$$

Method 2

Power dissipated is given by,

$$P = I^2 \times R = I \times I \times R \quad \dots\dots[1]$$

From Ohm's law $I = \frac{V}{R}$

substituting this in equation $\dots\dots[1]$

$$\text{Power dissipated} = I \times I \times R = \frac{V}{R} \times I \times R = V \times I$$

This formula for the power dissipated is advantageous because, by measuring the total circuit current and the applied circuit voltage, power consumed by the circuit can be found out easily without knowing the value of the resistors.

Using this formula for the circuit as given in Fig 5,

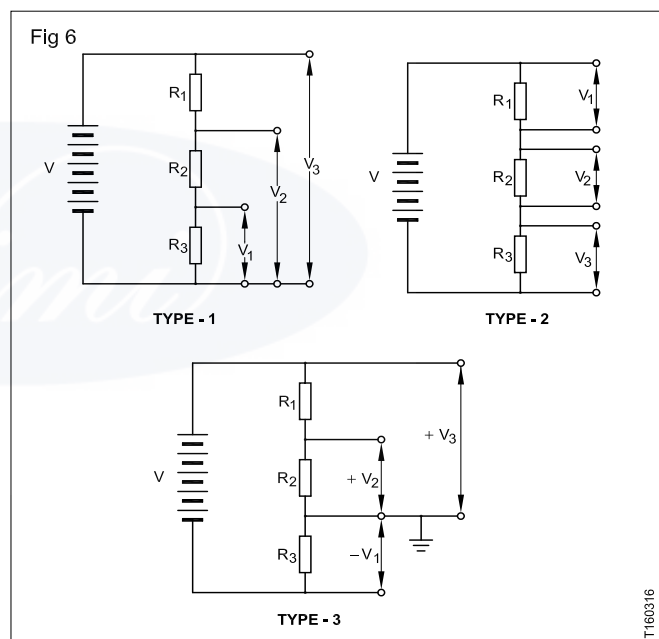
Power dissipated = $(6 \times 10^{-3}) \text{ amps} \times 12 \text{ volts} = 72 \text{ milli-Watts.}$

The power dissipated by the circuit found using method 2 is same as that calculated using method 1.

Try to derive alternative formulas for calculating power.

USE OF RESISTORS IN SERIES AS A VOLTAGE DIVIDER

Given a single voltage source, several circuits requiring smaller voltages can be obtained using a set of resistors connected in series. This is called a voltage divider. Some simple voltage divider circuits are shown in Fig 6.



Designing voltage divider circuits using resistors

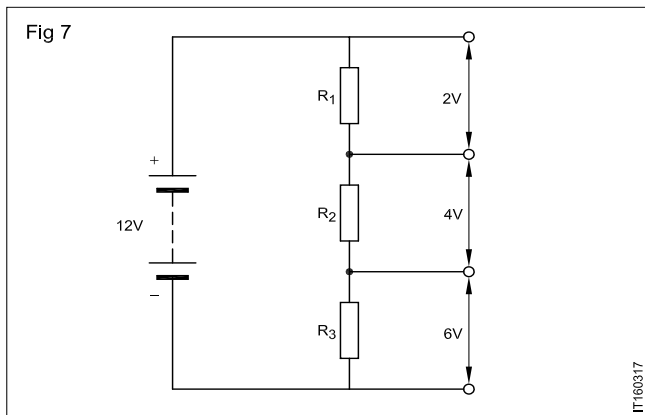
Example : Given a 12 volts DC power supply, design a voltage divider using resistors to obtain three voltages of 2 volts, 4 volts and 6 volts.

Solution :

Draw the divider diagram as shown in Fig 7. Assume the voltage across resistors as given in Fig 7.

Let the resistor values be R_1, R_2 & R_3 .

Assume a value of 1KW will have a drop of 1 volt across it.



Therefore to obtain 2 volts we need $R_1 = 2K$ ohms.

Similarly for 4 volts $R_2 = 4K$ ohms

for 6 volts $R_3 = 6K$ ohms

Easier method [Ratio method]

Find the ratio of the three different voltages required:

$$2V:4V:6V = 1:2:3.$$

Add up the ratios $1+2+3 = 6$.

Divide 12 volts by 6. $12/6 = 2$. Fix 2 ohms or 2 K ohms or 2 M ohms as the resistance value of R_1 required to drop 2 volts.

Let us fix $R_1 = 2K$ ohms for the present (reasons in later lessons).

Multiply 2 K ohms by the next ratio value i.e 2.

$2 \times 2K$ ohms = 4 K ohms. This is the value required to drop 4 volts.

Multiply 2 K ohms by the next ratio value i.e 3.

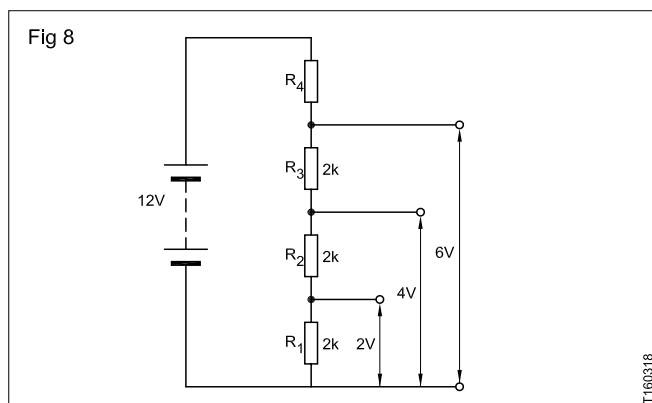
$2 \times 3K$ ohms = 6 K ohms. This is the value required to drop 6 volts.

Values of the resistors of the voltage divider should, therefore, be, $R_1 = 2K$ ohms, $R_2 = 4K$ ohms and $R_3 = 6K$ ohms.

The voltage divider circuit designed above can also be designed in a different way as shown in Fig 8.

Discuss merits and demerits of the above two designs.

Refer reference books listed at the end of this book for other methods of designing voltage dividers.

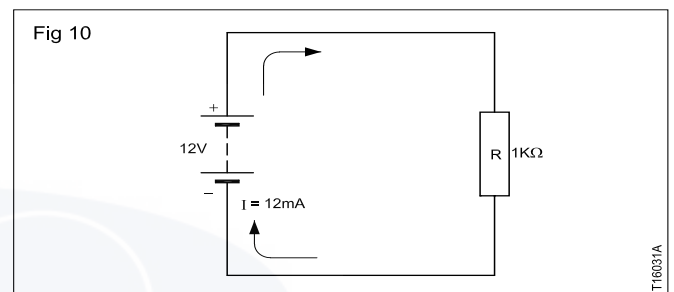
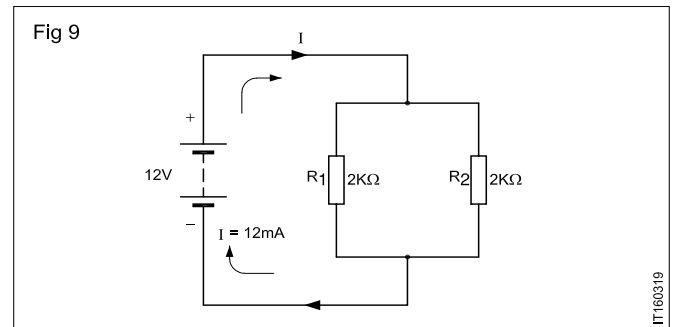


RESISTORS IN PARALLEL

Compare the circuits at Fig 9 and Fig 10. Both the circuits are carrying the same current although Fig 9 is having large resistor values than Fig 10.

This means that, the two resistors connected in parallel as shown in Fig 9 has effective or total resistance equal to that of a single resistor that is shown in Fig 10.

Analysing both circuits, it can be concluded that;

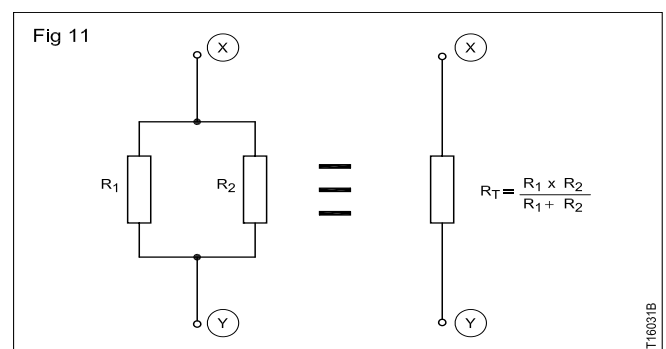


when resistors are connected in parallel as shown in Fig 11, the effective resistance (R_T) across the terminals x & y will be equal to,

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2}$$

or

$$R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$



Similarly, if 3 resistors are connected in parallel, the effective resistance R_T will be,

$$R_T = \frac{R_1 \times R_2 \times R_3}{R_1 R_2 + R_2 R_3 + R_1 R_3}$$

As a thumb rule When resistors are connected in parallel, the effective resistance will always be less than the least resistance value of the resistor in the parallel connection.

Example: If two resistors of $1K\Omega$ each are connected in parallel, from the thumb rule, the effective resistance will be less than $1K\Omega$.

Verification:

$$\text{From the formula derived above, } R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$

$$= \frac{1K \times 1K}{1K + 1K} = 0.5K\Omega$$

Example: If two resistors of $2K\Omega$ and 100Ω each are connected in parallel, from the thumb rule, the effective resistance will be less than 100Ω .

Verification

$$\text{From the formula derived above, } R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$

$$= \frac{2K\Omega \times 100\Omega}{2K + 100\Omega} = 95.24\Omega$$

Example: In circuit in Fig 11, if the value of the resistors were,

$R_1 = 10 \text{ ohms}$, $R_2 = 1 K \text{ ohm}$, what is the effective resistance?

Solution :

$$\text{Using the formula } R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$

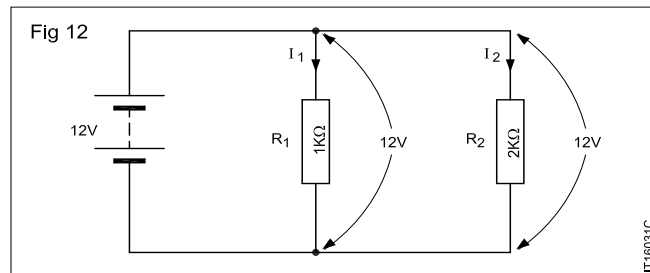
$$R_T = \frac{(10) \times (1 \times 10^3)}{(10) + (1 \times 10^3)} = \frac{10000}{1010} = 9.9 \text{ Ohms.}$$

Effective resistance of 10 ohms & $1 K\text{ohm}$ in parallel is 9.9 ohms or approximately 10 ohms . From this, it can be understood when a very small value resistor is connected in parallel with a very large resistor, the effective resistance of the parallel connection will be approximately equal to the values of the very small resistance itself.

The effective value of resistance becoming lower than the lowest value resistor in the parallel connection is called the **LOADING** effect. This effect is very important in electronic circuits. Effect of loading is discussed in detail in further lessons.

POWER DISSIPATION IN PARALLEL RESISTIVE CIRCUITS

When resistors are connected across a voltage source in parallel as in Fig 12, the voltage across each resistor will be the same and equal to the source/battery voltage.



However the current through the resistors (I_1 & I_2) depends on the value of the independent resistors. Hence, power dissipated/ consumed by each resistor can be calculated separately using the formula $P = I^2 R$ or $P = VI$. The sum of the power consumed by each resistors will be the total power consumed by the circuit.

Total power consumed can also be found if the total circuit current and the effective resistance of the circuit are found.

Example: Calculate the power consumed by the circuit in Fig 12.

Solution :

Method 1:

$$\text{Power consumed by the circuit} = (I_1^2 \times R_1) + (I_2^2 \times R_2)$$

$$I_1 = V/R_1 = 12/1K = 12 \text{ mA. } I_2 = V/R_2 = 12/2K = 6 \text{ mA.}$$

$$\text{Power dissipated by } R_1 = (12 \times 12 \times 10^{-6}) \times (1 \times 10^3) = 144 \text{ mWatts}$$

$$\text{Power dissipated by } R_2 = 6 \times 6 \times 10^{-6} \times 2 \times 10^3 = 72 \text{ mWatts}$$

Total power consumed by the circuit is

$$= 144 + 72 \text{ milli watts} = 216 \text{ milli-watts.}$$

Method 2:

$$\text{Total circuit resistance } R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$

$$= \frac{(1 \times 10^3) \times (2 \times 10^3)}{1 \times 10^3 + 2 \times 10^3} = 666.7\Omega$$

$$\text{Circuit current } I_t = I_1 + I_2$$

$$= \frac{12V}{666.7\Omega} = 0.018 \text{ Amps}$$

$$\text{Power consumed by the circuit} = I_t^2 \times R_T$$

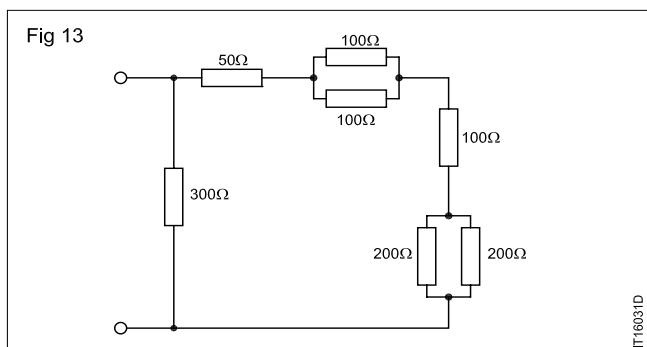
$$= (0.018)^2 \times (666.7)$$

$$= 0.216 \text{ watts}$$

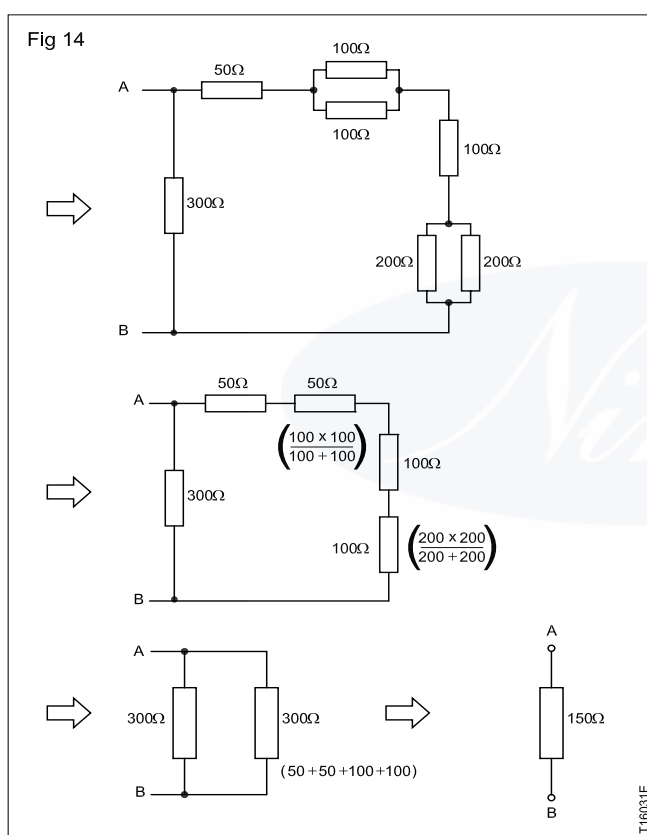
$$= 216 \text{ mwatts}$$

SERIES-PARALLEL RESISTOR CIRCUITS

As shown in Fig 13, when several resistors are connected, in which some resistors are in series and some in parallel, then, such a circuit is called a series-parallel circuit.



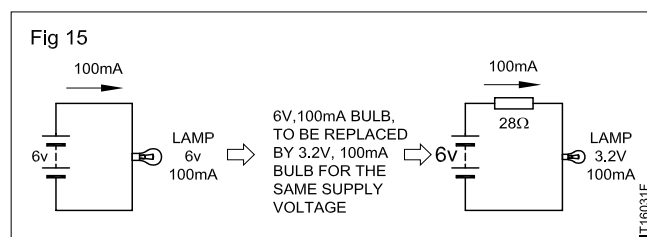
To find the effective resistance of such series-parallel circuits, follow the steps illustrated in Fig 14.



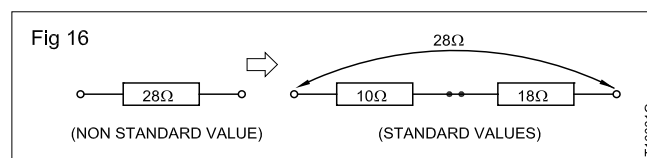
Once a complex series parallel circuit has been simplified and total circuit resistance is found, it is easy to calculate the circuit current and power consumed by the circuit. Using Ohm's law, it is also easy to calculate the voltage drop and power dissipated by individual resistors.

Some applications of series, parallel and series-parallel connection of resistors in circuit.

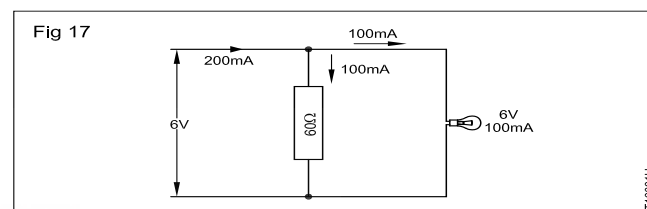
1 Series resistor to drop excess voltage.



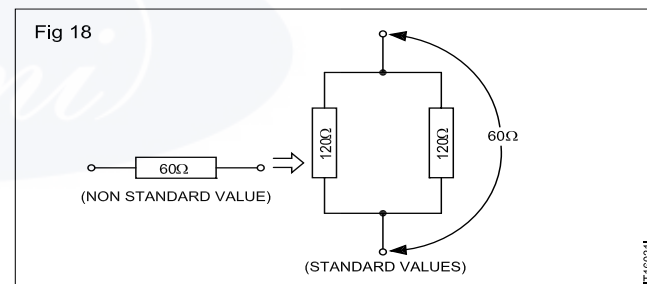
3 Parallel resistor used to divert excess circuit current.



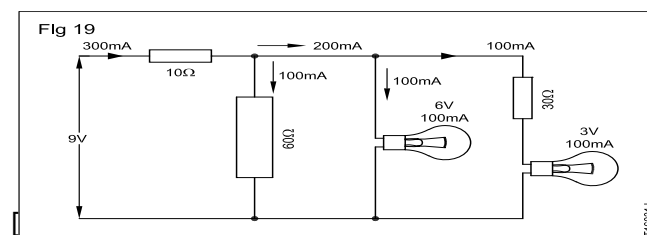
2 Series resistor to obtain non-standard resistor values.



4 Parallel resistor to obtain non-standard resistor values.



5 Resistors in series, parallel to drop voltage and divert current.



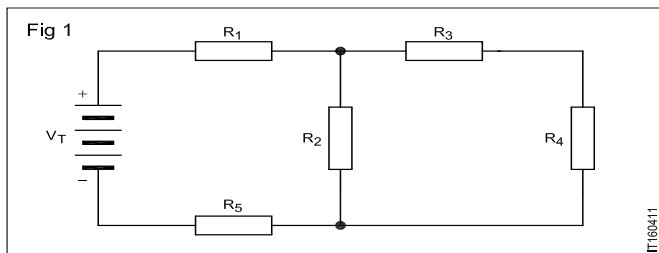
H.I.Discus the circuit given at Figures 15 to 19.]

Kirchhoff's Laws

Objectives: At the end of this lesson you shall be able to

- state Kirchhoff's current law
- write Kirchhoff's current equation for resistive circuits
- state Kirchhoff's voltage law
- write Kirchhoff's loop equation for resistive circuits
- calculate voltage and currents at different points in a resistive circuit.

When a circuit consists of several resistors in a complex series - parallel arrangement as in Fig 1, it is difficult to calculate the currents and voltages in the circuit using Ohm's law.

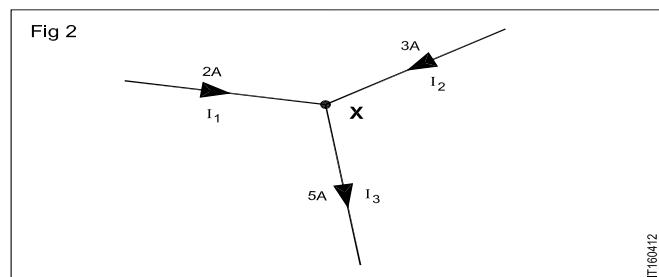


To find current and voltage drops in a complex series - parallel circuit, an easy method was found by a German physicist GUSTAV R. KIRCHHOFF. He formulated two basic laws called,

1. Kirchhoff's Current law
2. Kirchhoff's Voltage law.

1 KIRCHHOFF'S CURRENT LAW:

This law is illustrated in Fig 2.



KIRCHHOFF'S Current law states that The sum of currents entering any point in a circuit is equal to the sum of currents leaving that point.

In Fig 2 currents I_1 and I_2 are entering a point X. Current I_3 is leaving the point X.

From Kirchhoff's current law, $I_1 + I_2 = I_3$ [1]

This equation can also be written as,

$$I_1 + I_2 - I_3 = 0 \text{[2]}$$

From equation 2, Kirchhoff's current can also be stated as The algebraic sum of currents entering and leaving any point in a circuit must be equal to zero .

To determine the algebraic sign of currents,

- consider all currents going into a point as *positive* and all currents going away from that point as *negative*.

In Fig 2, I_1 & I_2 will have positive sign as they are going into point x whereas I_3 will have negative sign as it is going out of the point X.

Hence we can also write the Kirchhoff's Current equation as,

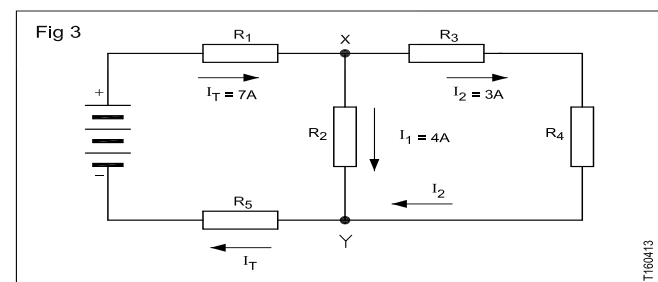
At point X,

$$(+I_1) + (+I_2) + (-I_3) = 0$$

Simplifying, $I_1 + I_2 - I_3 = 0$

Substituting current values given in Fig 2,
2Amps + 3Amps - 5Amps = 0.

For the circuit shown in Fig 3, Kirchhoff's Current equation at nodes X and Y can be written as follows:



At node X $I_T - I_1 - I_2 = 0$

$$7A - 4A - 3A = 0.$$

At node Y $I_1 + I_2 - I_T = 0$

$$4A + 3A - 7A = 0.$$

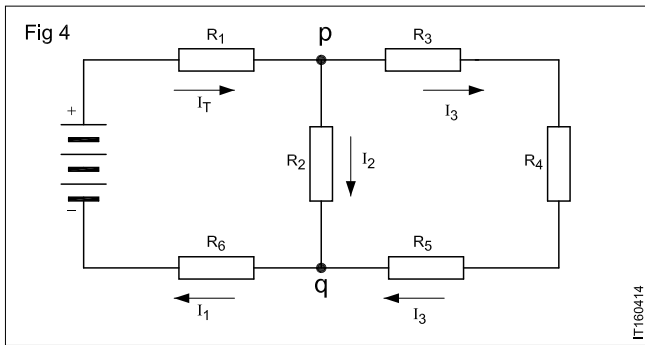
Example : Write the current equations at nodes p and q in the circuit at Fig 4.

At node P $I_T - I_2 - I_3 = 0$

or $I_T = I_2 + I_3$

At node Q $I_2 + I_3 - I_1 = 0$

or $I_2 + I_3 = I_1$



KIRCHHOFF'S VOLTAGE LAW

In the circuit shown at Fig 5A, consider the two closed paths a-c-d-b-a and a-e-f-b-a as shown in Fig 5b. These closed paths are called as *loops*. Each closed path has several resistors and there will be a voltage drop across each resistor. KIRCHHOFF's voltage law states that The algebraic sum of voltages around any closed path is zero.

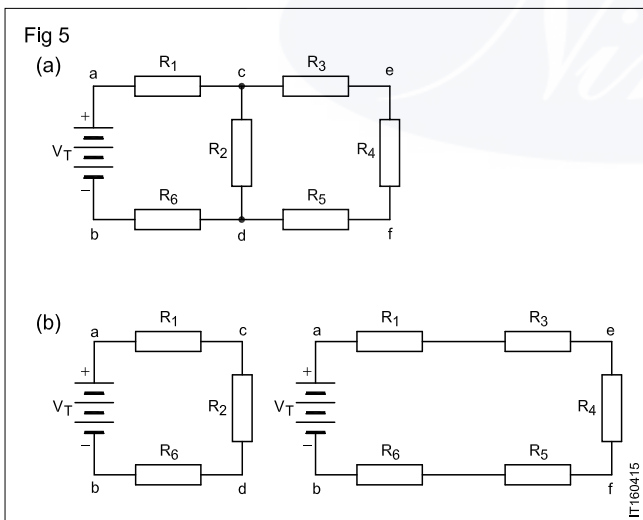
To find the algebraic sum of around a closed path,

- start from any point, go around the path and come back to the same point from where you started.

Example: Referring to Fig 5b, the method of going through a closed path is,

- start from point **a**, go through points **c,d,b** and return to point **a**.

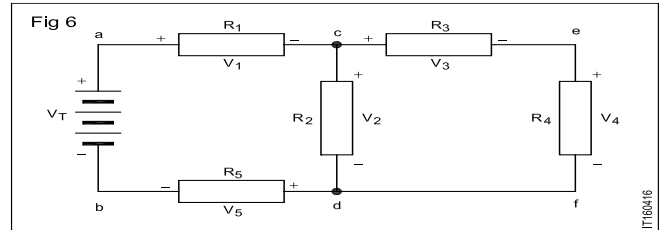
To determine the signs for voltage drop across the resistors in Fig 5b,



- mark the polarity of each voltage, based on the polarity of source voltage as shown in Fig 6
- go around the path and give +ve sign for the voltage whose PLUS terminal is reached first or give -ve sign for the voltage whose MINUS terminal is reached first.

Write the loop equation considering the voltage sources also.

To write the loop equation for the closed path **a-c-d-b-a** of Fig 6, proceed as follows:



Taking clockwise direction for going through the loop, start from point **a** of Fig 6. Go through the chosen loop **a-c-d-b-a** and write down the voltage drop across the resistors including their signs and equate it to zero as given below;

$$+V_1 + V_2 + V_5 - V_T = 0 \quad \dots\dots[1]$$

Rewrite the equation as,

$$+V_1 + V_2 + V_5 = V_T.$$

Similarly for the closed path **a-e-f-b-a**,

considering clockwise direction, start from point **a** of Fig 6. Go through the chosen loop **a-e-f-b-a** and write down the voltage drop across the resistors including their signs and equate it to zero as given below;

$$+V_1 + V_3 + V_4 + V_5 - V_T = 0 \quad \dots\dots[2]$$

Rewriting the equation,

$$+V_1 + V_2 + V_3 + V_4 = V_T.$$

Equations [1] & [2] above state that; In any closed loop, the sum of voltage drops across resistors is equal to the applied voltage. This can be written as:

$$S V_d = V_T,$$

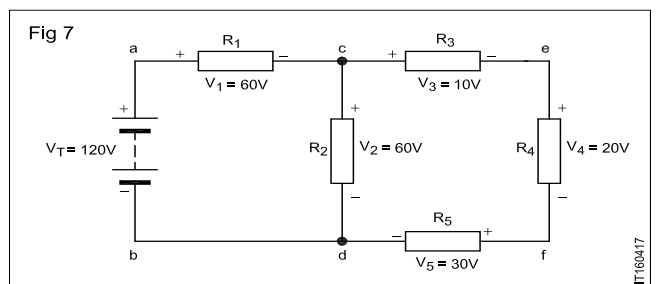
where, $S V_d$ is the sum of voltage drops across resistors V_T is the applied voltage.

$$\text{For the loop a-c-d-b-a,} \quad +V_1 + V_2 - V_T = 0$$

or

$$V_1 + V_2 = V_T$$

Example : Write the loop equations for the circuit given at Fig 7.



Verification

$$60 + 60 = 120$$

For the loop a-e-f-b-a

or

$$+V_1 + V_3 + V_4 + V_5 - V_T = 0$$

$$V_1 + V_3 + V_4 + V_5 = V_T$$

Verification

$$60 + 10 + 20 + 30 = 120$$

For the loop c-e-f-d-c

$$+V_3 + V_4 + V_5 - V_2 = 0$$

$$\text{or } V_3 + V_4 + V_5 = V_2$$

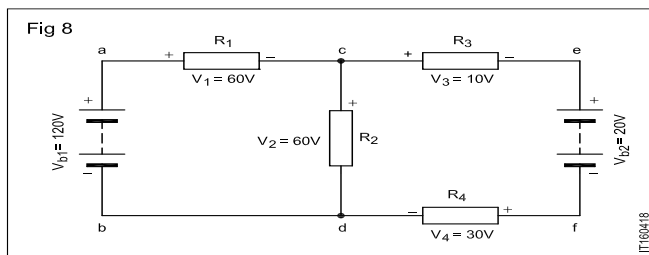
Verification

$$10 + 20 + 30 = 60$$

Circuit with more than one voltage source

Kirchhoff's voltage law is applicable even when, there are more than one voltage source in a circuit. The method of writing loop equations remains the same.

Example: Write the loop equations for the circuit shown in Fig 8.



For the loop a-c-d-b-a,

$$+V_1 + V_2 - V_{b1} = 0$$

$$V_1 + V_2 = V_{b1}$$

$$60 + 60 = 120$$

For the loop a-e-f-b-a,

$$+V_1 + V_3 + V_{b2} + V_4 - V_{b1} = 0$$

$$V_1 + V_3 + V_{b2} + V_4 = V_{b1}$$

$$60 + 10 + 20 + 30 = 120$$

For the loop c-e-f-d-c,

$$+V_3 + V_{b2} + V_4 - V_2 = 0$$

$$V_3 + V_{b2} + V_4 = V_2$$

$$10 + 20 + 30 = 60$$

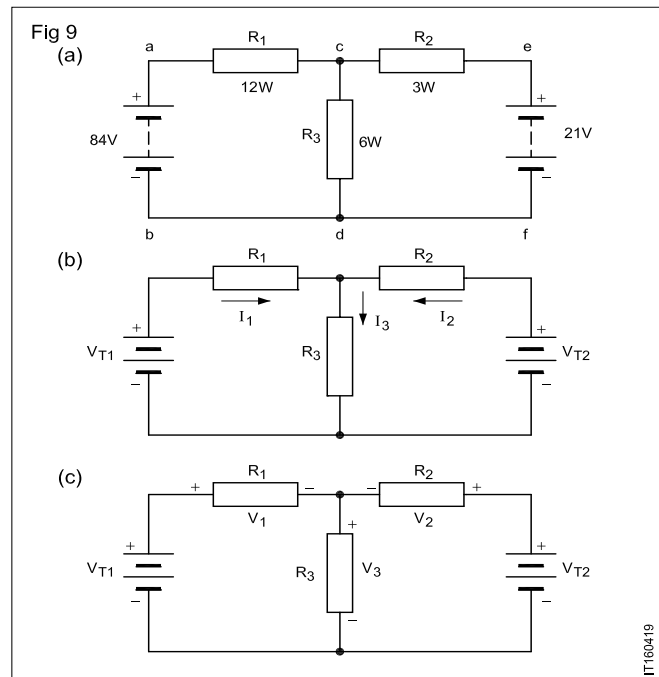
To find current and voltages in a circuit using Kirchhoff's law

Example: In the circuit at Fig 9a, find the branch currents and voltage drop across resistors R_1 , R_2 and R_3 .

Solution :

Step 1: Assume the current flow is from the +ve of the battery to the -ve of the battery (conventional current direction) and write direction of current as shown in Fig 9b.

Step 2: Mark the polarity of voltage across each resistor as shown in Fig 9c consistent with the assumed direction of the current.



Step 3: Write the loop equations which cover the two battery sources as given below;

For the loop a-c-d-b-a,

$$+V_1 + V_3 - V_{T1} = 0$$

$$\text{or } V_1 + V_3 = V_{T1}$$

$$V_1 + V_3 = 84 \quad \dots\dots[1]$$

For the loop c-e-f-d-c,

$$-V_2 + V_{T2} - V_3 = 0$$

$$\text{or } V_2 + V_3 = V_{T2}$$

$$V_2 + V_3 = 21 \quad \dots\dots[2]$$

Using the known values of R_1 , R_2 & R_3 , write I_R drops, V_1 , V_2 and V_3 .

$$V_1 = I_1 \cdot R_1 = I_1 \times 12 = 12 \cdot I_1 \quad \dots\dots[3]$$

$$V_2 = I_2 \cdot R_2 = I_2 \times 3 = 3 \cdot I_2 \quad \dots\dots[4]$$

$$V_3 = I_3 \cdot R_3 = (I_1 + I_2) \times 6 = 6 \cdot I_1 + 6 \cdot I_2 \quad \dots\dots[5]$$

Substitute equations 3, 4 and 5 in the loop equations 1 and 2, we have,

Substituting for equation 1,

$$V_1 + V_3 = 84$$

$$\text{or } (12 \cdot I_1) + (6 \cdot I_1 + 6 \cdot I_2) = 84$$

$$18 \cdot I_1 + 6 \cdot I_2 = 84$$

(Dividing RHS & LHS by 2)

$$9 \cdot I_1 + 3 \cdot I_2 = 42 \quad \dots\dots[6]$$

Substituting for equation 2,

$$\begin{aligned} V_2 + V_3 &= 21 \\ (3.I_2) + (6.I_1 + 6.I_2) &= 21 \\ 6.I_1 + 9.I_2 &= 21 \end{aligned}$$

(Divide RHS & LHS by 3)

$$2.I_1 + 3.I_2 = 7 \quad \text{.....[7]}$$

Solving equations ..[6] and ..[7],

$$9.I_1 + 3.I_2 = 42$$

$$2.I_1 + 3.I_2 = 7$$

$$7.I_1 = 35$$

$$\text{or } I_1 = \frac{35}{7} = 5 \text{ amps.}$$

Substituting value of I_1 in equation ...[7]

$$2 \times 5 + 3.I_2 = 7$$

$$\text{Therefore } I_2 = -\frac{3}{1} = -1 \text{ amp.}$$

The Minus sign indicate that, the assumed direction of I_2 in Fig 9b was opposite to that of the actual direction of the current. However the numerical value of I_2 remains unchanged.

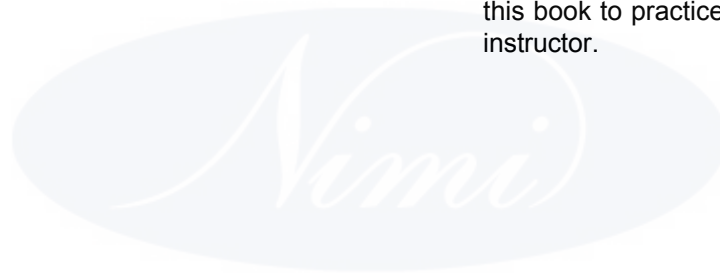
From known value of I_1 , I_2 and $I_3 = (I_1 + I_2)$, find the voltage drop across resistors as given below;

$$V_1 = R_1.I_1 = 12 \times 5 = 60 \text{ volts}$$

$$V_2 = R_2.I_2 = 3 \times 1 = 3 \text{ volts}$$

$$V_3 = R_3.I_3 = R_3.(I_1 - I_2) = 6 \times (5 - 1) = 24 \text{ volts.}$$

- 1 The method used for finding branch currents and voltage drops in the above discussion is called **Branch current method**.
- 2 To solve similar problems, other methods known as, **Node voltage method, Mesh current method** can also be used. Refer reference books listed at the end of this book to practice these methods or consult your instructor.



Electrical power

Objectives: At the end of this lesson you shall be able to

- state the formula for calculating electrical power
- use the formula in conjunction with Ohm's law to describe the direct and indirect methods of measuring electrical power
- state the relation between mechanical and electrical power
- define the termse related to work & power.

Electrical power: The rate at which electrical energy is delivered to a load is called ELECTRICAL POWER. The unit of measurement of electric power is the WATT, and the symbol for power is P. In electrical circuits the power is equal to the current, multiplied by the voltage or $P = I \times E$. Therefore, one watt of power is the result of one ampere of current driven by a one volt force through a circuit.

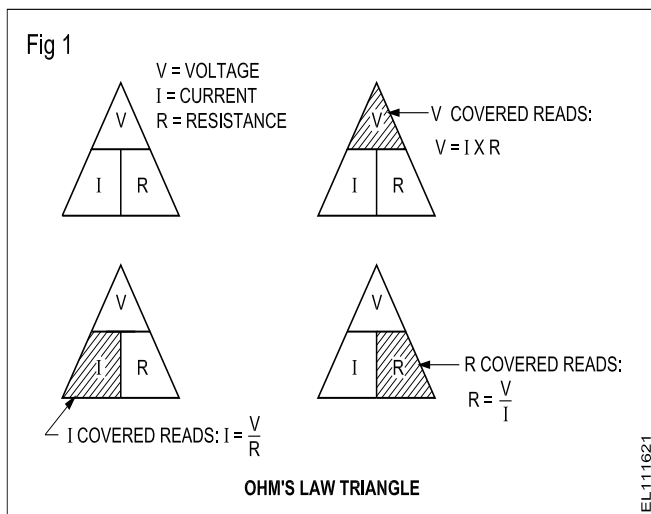
Example : A heater takes 4 amps at 250 volts. How much power does it consume?

$$\begin{aligned} P &= E \times I \text{ watts} \\ &= 250 \times 4 = 1000 \text{ watts.} \\ &= 1 \text{ kilowatt} \end{aligned}$$

By simple algebra, we can write the 'Power Law' or 'Watts Law' in three ways as we did for Ohm's law.

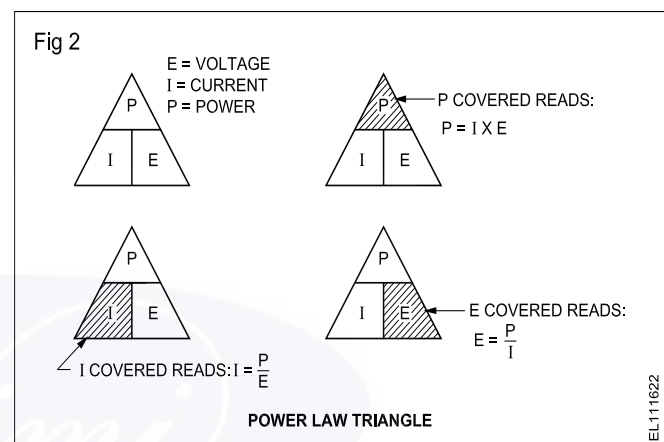
$$\begin{aligned} P &= I \times E \\ I &= \frac{P}{E} \\ E &= \frac{P}{I} \end{aligned}$$

You know how to apply the memory device of Ohms law. To recapitulate, it is shown in Fig 1.

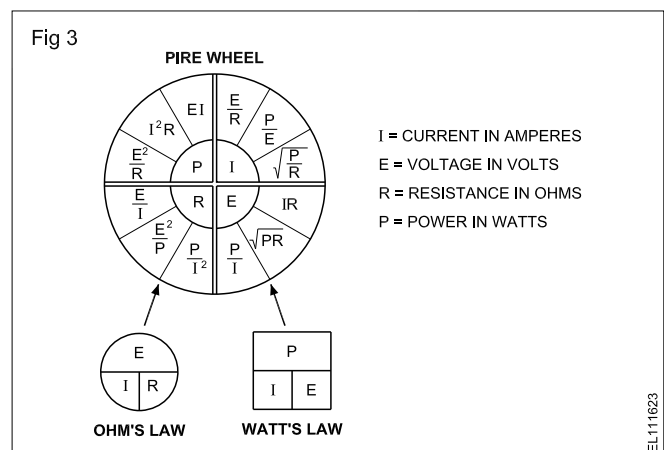


However, power in an electrical circuit could be measured directly with a waltmeter.

The three power law formulas can be easily recalled by arranging the three quantities within a triangle as shown in Fig 2.



The formulae (or equations) to solve for unknown voltage, current, resistance or power can be obtained by combining Ohm's law and Power law. This is shown in Fig 3.



Question: A certain appliance has a resistance of 100 ohms is connected to 200V supply. What is the power?

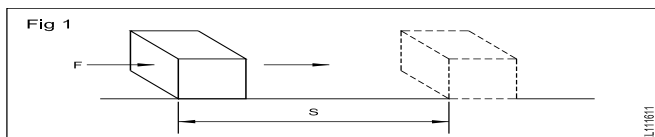
In your exercise book work out the answer to the problems using the pire wheel.

- 1 $P = 500W$, $E = 250$, $I = ?$
- 2 $P = 10W$, $I = 1amp$, $R = ?$
- 3 $I = 0.001amp$, $R = 1Megohm$, $P = ?$
- 4 $I = 100mA$, $R = 100ohms$, $P = ?$
- 5 $P = 10W$, $I = 2amp$, $E = ?$

Measure AC & DC power

Work: This may be defined as the overcoming of resistance or obstruction through a distance. The unit of measurement of work is joule which is equal to 0.102 kilogram-meter.

If a force of 'F' moves a body through a distance 'S' (Fig 1) in its direction of application, then,



$$\text{work done } W = F \times S$$

$$\text{If } F = 1N; S = 1m \text{ then}$$

$$\text{work done} = 1 \text{ Nm or joule.}$$

Hence one joule is the work done when a force of 1 N moves a body through a distance of 1m in the direction of its application.

Force: The absolute unit of force is 'newton' (N). One newton is that force which accelerates a mass of one kilogram at 1m per s^2 . This definition is obtained from Newton's Law of Motion.

$$\text{That is } F = ma.$$

Power: Machines were invented to assist in doing work. The amount of work a machine can do in a specific time determines its power.

Power is the measurement of the rate of doing work, and is the product of force and the distance moved in 1 second.

If a force of F newton moves a body with a velocity of v m/s then

$$\text{power} = F \times v \text{ watt.}$$

Horsepower: Horsepower (HP) is the common unit of measurement of mechanical power

$$1 \text{ hp} = 75 \text{ mkg/s (Fig 2)}$$

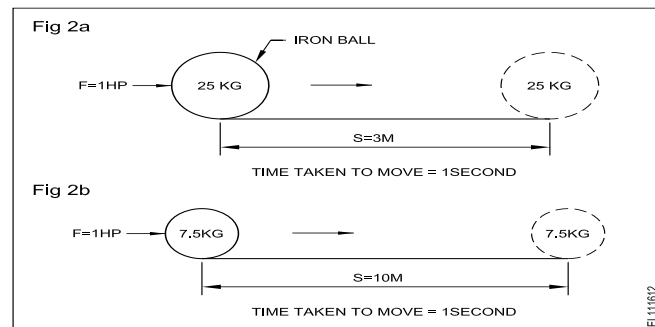
$$= 75 \times 9.81$$

$$= 735.5 \text{ Nm/s or watt}$$

$$(1 \text{ kg.weight} = 9.81 \text{ N.})$$

As shown in Fig 2a, 2b and 3, the product of force x distance divided by time gives the horse power rating of the power required to do the job.

(Refer Fig 2a)

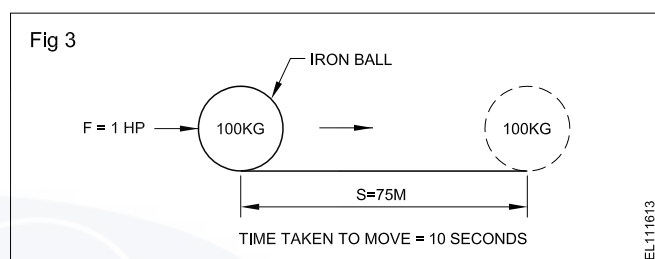


$$HP = \frac{\text{Work done}}{\text{Time required in sec.}} = \frac{25 \times 3}{1 \text{ sec.}} = 75 \text{ mkg / sec.}$$

Refer Fig 2b.

$$HP = \frac{\text{Work done}}{\text{Time required in sec.}} = \frac{7.5 \times 10}{1 \text{ sec.}} = 75 \text{ mkg / sec.}$$

Find the hp as shown in Fig 3.



Example 1 : A force of 60 N is applied to a body to move it at a uniform velocity through a distance of 20m in 8 seconds in the direction of the force. Calculate the value of the power.

$$\text{Work done} = 60 \text{ [N]} \times 20\text{[m]} = 1200 \text{ J}$$

$$\begin{aligned} \text{Therefore, power} &= \text{work done per second} \\ &= 1200 \text{ [J]} / 8\text{[s]} \\ &= 150 \text{ W.} \end{aligned}$$

Example 2: Calculate the power required to lift a mass of 300 kg at a constant speed through a vertical distance of 200m in 4 minutes.

$$\begin{aligned} \text{Force required to lift load} &= 300 \text{ [kg]} \times 9.81 \text{ [N/kg]} \\ &= 2943 \text{ N} \end{aligned}$$

$$\begin{aligned} \text{Work done in 4 min} &= 2943 \text{ [N]} \times 200 \text{ [m]} \\ &= 588 \text{ 600J} \end{aligned}$$

$$\text{Therefore, power} = \frac{588 \text{ 600J}}{(4 \times 60)(s)}$$

Wattmeter

Objectives: At the end of this lesson you shall be able to

- state the advantages of measuring power directly
- state the types of single phase wattmeters
- explain the construction and working of the induction type single phase wattmeter
- choose the correct method of connecting pressure and current coils of the wattmeter in a circuit to reduce measurement errors.

Power in a single phase AC circuit can be calculated by using an ammeter, a voltmeter and a power factor meter with the help of the formula

Power in a single phase circuit = $E I \cos \phi$ watts.

In the same way power in a 3-phase balanced circuit can be measured by using one ammeter, one voltmeter and one power factor meter with the help of the formula

Power in a balanced 3 - phase circuit =

$$\text{or, } 3E_P I_P \cos \phi$$

where $E_P I_P$ are the phase values and $\sqrt{3} E_L I_L \cos \phi$

$E_L I_L$ are the line values

This indirect method of measuring power in an AC circuit has the following disadvantages.

- Low accuracy due to a number of meters.
- Reading errors due to a number of meter readings.
- Cumbersome connections.
- Involves calculation every time the load changes and hence not suitable for changing loads.

To get an on the spot true power reading, a wattmeter is used. The power dissipated in the circuit can be read directly from the scale of the meter. The wattmeter takes the power factor of the circuit into account and always indicates the true power.

Types of wattmeters

There are three types of wattmeters in use as stated below.

- Dynamometer wattmeter
- Induction wattmeter
- Electrostatic wattmeter

Among the three, the electrostatic type is very rarely used. Information given here is for the other two types only.

Dynamometer type, single phase wattmeter: This type is commonly used as a wattmeter.

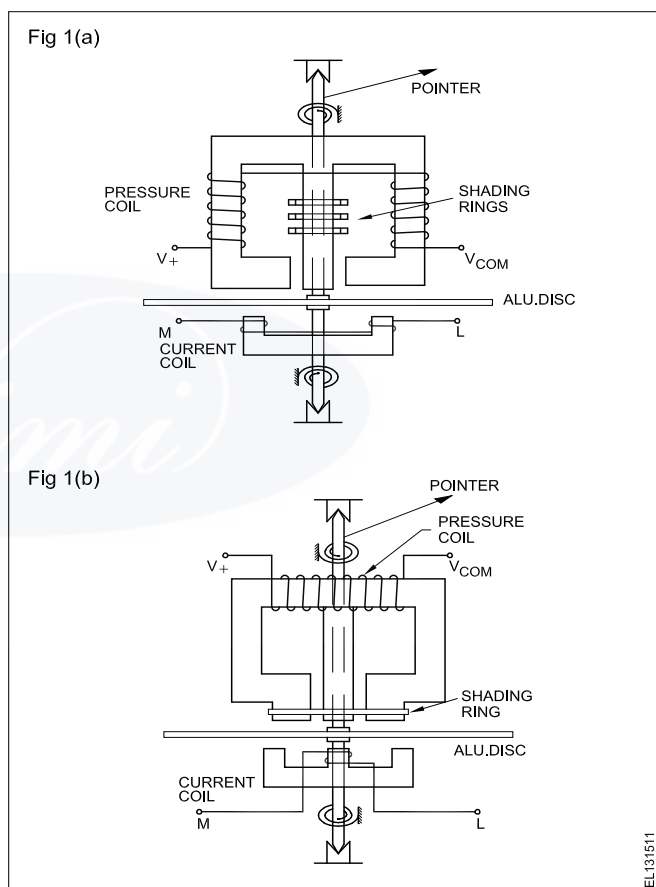
Induction type single phase wattmeter: This type of wattmeters could be used only in AC circuits whereas a dynamometer type wattmeter could be used in both AC and DC circuits.

Induction type wattmeters are useful only when the supply voltage and frequency are almost constant.

Construction: Induction wattmeters having two different types of magnetic cores are shown in Figs 1a and 1b.

Both the types have one pressure coil magnet and one current coil magnet. The pressure coil carries a current proportional to the voltage whereas the current coil carries the load current.

A thin aluminium disc is mounted on a spindle in between the space of the magnets and its movement is controlled



by springs. The spindle carries a weightless pointer at one end.

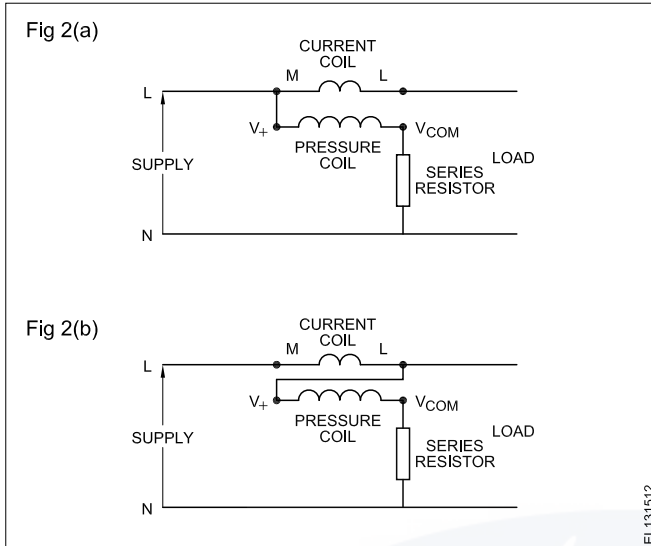
Working: The alternating magnetic fluxes produced by the pressure and current coils cut the aluminium disc and produce eddy currents in the disc. Due to the interaction between the fluxes and the eddy currents a deflecting .

Control springs attached to the two ends of the spindle control the deflection and the pointer shows the power in watts on a graduated scale.

Shaded rings provided in the pressure coil (shunt) magnet could be adjusted in order to cause the resultant flux in the magnet to lag in phase by exactly 90° behind the applied voltage.

Method of connecting wattmeter in single phase circuits - pressure coil connection to reduce erroneous measurement.

There are two ways of connecting the pressure coil of the wattmeter as shown in Fig 2.



Both the methods shown in Figs 2a & b need correction in power measurement due to the reasons stated below.

In the method of connection shown in Fig 2a, the pressure coil is connected on the 'supply' side of the current coil, and hence, the error in power measurement is due to the fact that the voltage applied to the voltage coil is higher than that of the load on account of the voltage drop in the current coil. As such the wattmeter measures the load power in addition to the power lost in the current coil.

On the other hand, in the method of connection shown in Fig 2b, the current coil carries the small current taken by the voltage coil, in addition to the load current, thereby introducing errors in power measurement. As such the wattmeter measures the load power in addition to the power lost in the pressure coil.

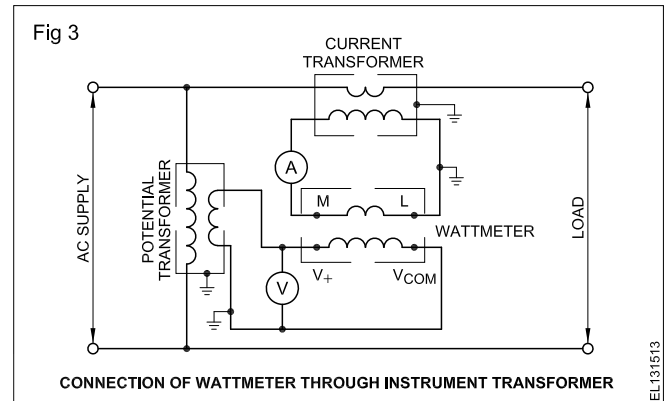
If the load current is small, the voltage drops in the current coil will be small, so that the method of connection, shown in Fig 2a, introduces a very small error and, hence, preferable.

On the other hand, if the load current is large the power lost in the pressure coil will be negligible when compared to the load power in the method of connection shown in Fig 2b, and, hence, a very small error is introduced resulting in the preference of this connection.

Connecting wattmeters through instrument transformer: Normally the rating of the current coil of the wattmeter is in the order of 1A, 5A, 10A, 20A, or 30 amps, whereas the pressure coil rating will be in the order of 110V, 220V, 300V, 500V. When the circuit in which the power to

be measured has a current more than 30 amps and voltages more than 500V it is customary to use current and potential transformers along with the wattmeter to measure power in the circuit.

In practice, a wattmeter having a current coil rating of 5 amps and a pressure coil rating of 110V could be used along with the instrument transformers of different ratios to measure power in a variety of circuits having high current and high voltages. Connection of a wattmeter along with the instrument transformers is shown in Fig 3 in which an ammeter and a voltmeter are also connected in the circuit of the instrument transformers.



The primary winding of the current transformer is connected in series with the load and the secondary winding is connected in series with the current coil of the wattmeter and ammeter.

The primary winding of the potential transformer is connected across the supply, and the secondary winding is connected in parallel with a voltmeter and the potential coil of the wattmeter. For ensuring safety, one of the terminals of secondary winding and the casing of the instrument transformer are earthed.

Assuming there is negligible ratio and phase angle error in the instrument transformers, the power measured by the wattmeter needs to be multiplied by the C.T. and P.T. ratios to obtain the actual power consumed in the circuit.

Hence we have

$$\text{Actual power consumed in the circuit} = \text{Ratio of CT} \times \text{Ratio of PT} \times \text{wattmeter reading.}$$

Example: A wattmeter whose current coil rating of 5 amps and the voltage coil rating of 110V is connected through an instrument transformer and reads 400 watts. The CT ratio is 20:1 and the PT ratio is 100:1. Calculate the

- voltage of the mains supply
- current in the main circuit
- actual power consumed by the circuit.

$$\begin{aligned} \text{a) Voltage of the main supply} &= \text{Pressure coil rating of wattmeter} \times \text{P.T. ratio} \\ &= 110 \times 100/1 = 11000\text{V} = 11 \text{ KV.} \end{aligned}$$

$$\begin{aligned} \text{b) Current in the main circuit} &= \text{Current coil rating of wattmeter} \times \text{C.T. ratio} \\ &= 5 \times 20/1 = 100 \text{ amps} \end{aligned}$$

c) The actual power = Ratio of C.T. x Ratio of P.T.
 consumed in the x Wattmeter reading
 main circuit

$$= 20/1 \times 100/1 \times 400$$

$$= 800000 \text{ watts} = 800\text{KWatts}$$

$$= 0.8 \text{ megawatts}$$

Alternatively we have

Actual power = Multiplying factor of the
 consumed in the instrument transformer
 circuit x wattmeter reading.

As such,

Multiplying factor = Ratio of CT x Ratio of PT
 instrument transformer.



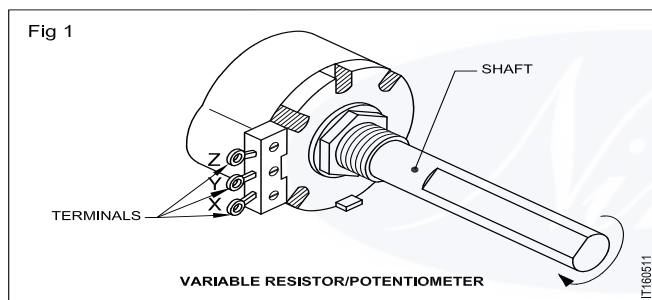
Special types of resistors

Objectives: At the end of this lesson you shall be able to

- state the different types of variable resistors based on the
 - (a) type of control
 - (b) resistive material used
 - (c) variation of resistance with respect to shaft movement
- state the important parts of a carbon potentiometer
- state the value of Lin and Log Potentiometers at different shaft positions
- state advantages of wire wound Potentiometers over Carbon Pots
- list a few special purpose Potentiometers
- list the general defects of potentiometers
- list the care to be taken while mounting Potentiometers
- list the specification to indicate while purchasing any Pot.

VARIABLE RESISTORS

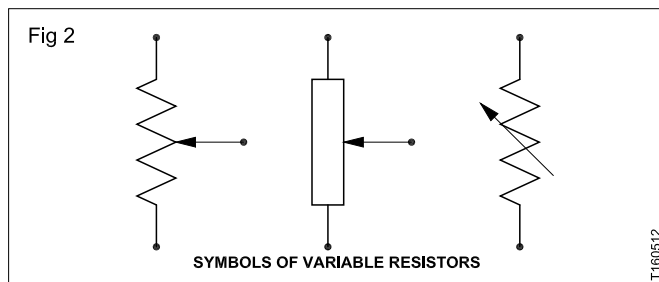
In fixed value resistors the value of resistance is fixed, at the time of manufacture, whereas, variable resistors are manufactured such that its resistance value can be changed by the user. A typical variable resistor, commonly known as potentiometer is shown in Fig 1.



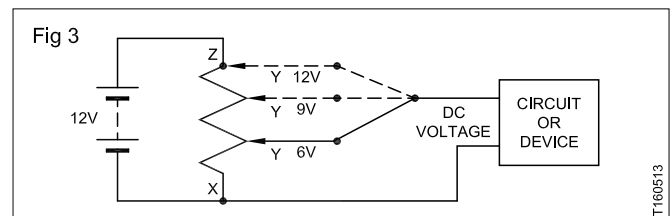
For example, in a variable resistor of 10 KW, it is possible to set the resistance of the variable resistor, to any value between 0 W to 10 K Ω .

Any variable resistor will have a minimum of three terminals. Referring to Fig 1, the maximum resistance exists across the end terminals X & Z. The resistance across terminals X & Y or across Z & Y can be varied by rotating the shaft.

Different symbols used to represent variable resistors are shown in Fig 2.



Variable resistors are also known as Potentiometers or Pots. The term potentiometer is used for variable resistors because, a variable resistor can be used to adjust the potential difference (voltage) applied to a device or a circuit as shown in Fig 3.

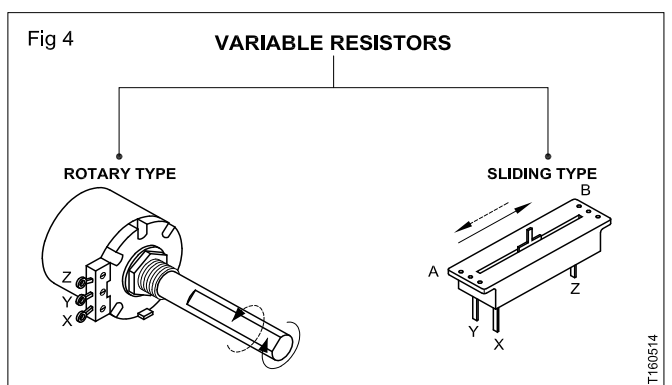


Resistance values of potentiometers are printed on their body either directly in ohms or in a coded form. The value printed always refers to the maximum value that potentiometer can have. Hence the minimum value to which it can be set is always 0 Ω .

Types of variable resistors

Variable resistors of different types are available in the market. These types can be categorized as follows:

1 Based on the type of control as shown in Fig 4.



In rotary type of variable resistors, the resistance across the terminals X-Y increases as the shaft is rotated anti-clock-wise whereas in sliding type, resistance across X-Y increases as the slider is moved from A to B.

2 Based on the resistive material used as shown in Fig 5.

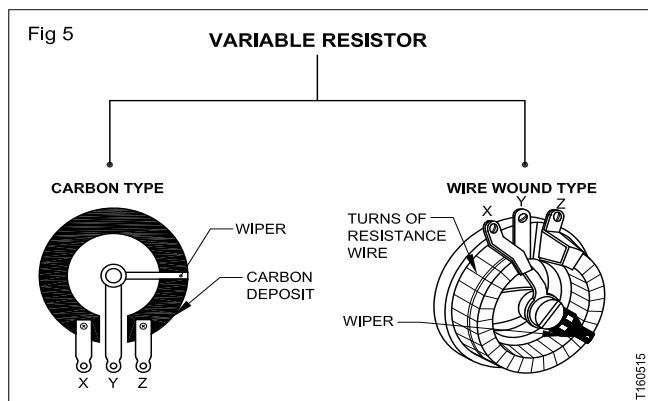
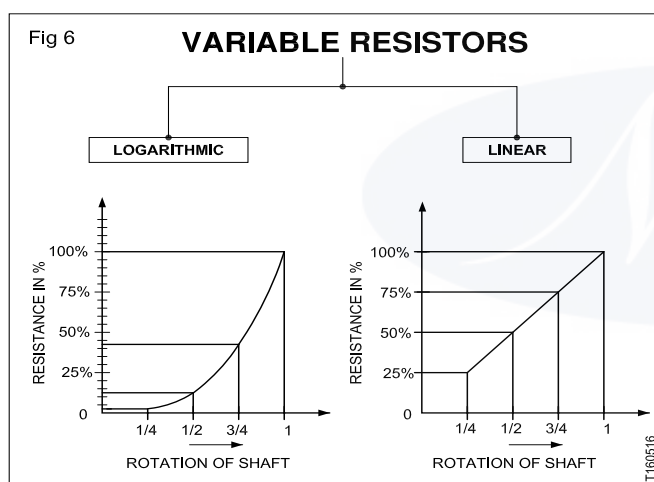


Fig 5 illustrates the brief constructional detail of carbon type and wire wound type variable resistors. In carbon type, the resistive material used is carbon or graphite. Wire wound type variable resistors use several turns of resistive wire such as tungsten or manganin.

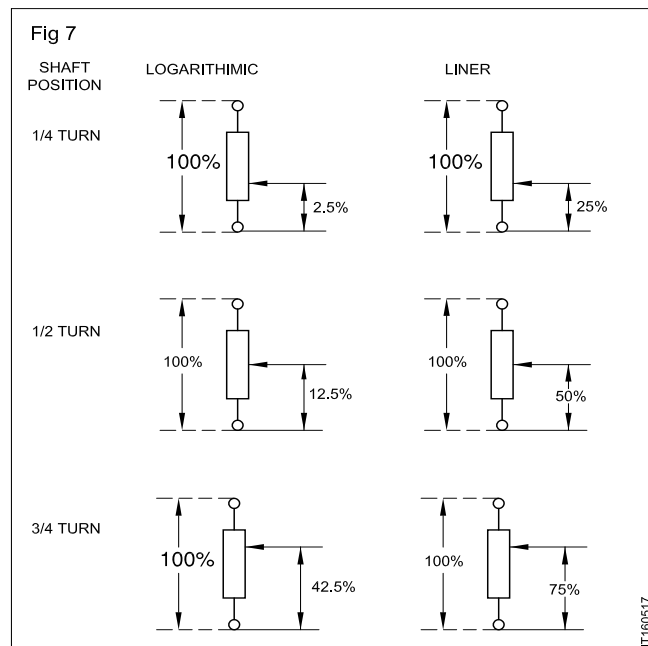
3 Based on variation of resistance with reference to shaft position (no. of turns) as shown in graph at Fig 6.



In linear variable resistors, the resistance across the resistor terminals increases or decreases linearly with respect to the shaft position as shown in the graph at Fig 6.

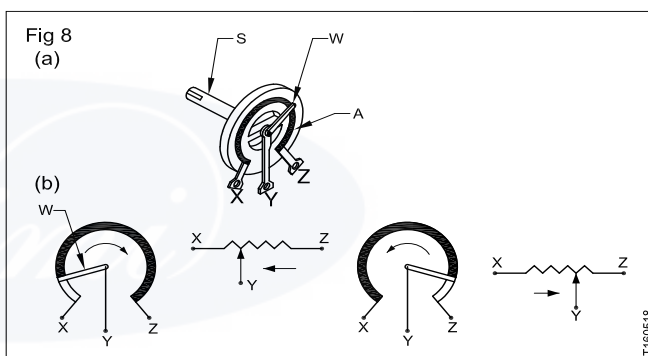
In logarithmic variable resistors, the resistance across the terminals increases or decreases logarithmically (as shown in graph at Fig 6) with respect to the shaft position.

Percentage variation of resistance at different shaft positions of the linear and logarithmic resistors is given in Fig 7 below;



Constructional details of variable resistors

Brief construction of a typical carbon type rotary shaft, variable resistor is shown in Fig 8.



The resistive material carbon or graphite with a suitable synthetic resin is deposited over an arc (A). The ends of the deposit are terminated to lugs (X and Z). The rotatable slider or wiper (W) made of a good conductor, makes contact with the deposited graphite. The other end of the wiper is terminated to the middle lug (Y). The wiper is mechanically coupled to a shaft (S). As the shaft is rotated, the wiper makes contact at different points on the carbon deposit and the resistance value across X & Y or Z and Y terminals increases or decreases.

In the case of wire-wound type variable resistors, instead of carbon/graphite deposit, suitable number of turns of resistive wire will be wound on an insulating former as shown in Fig 5.

Variable resistors are commonly used in electronic circuits. For ease of mounting and controlling, variable resistors are available in different shapes, sizes. Depending on their type and use they are called by different names such as, Potentiometer or POTs, Presets, Rheostats and so.. on. Refer to Chart 1 given at the end of this lesson for the most commonly used types of variable resistors.

Single turn carbon potentiometers

These are the most common of all types of variable resistors. These are generally referred to as single turn POTs. These are available in both rotary and sliding types as shown in Chart 1.

These pots are most commonly used as volume controls of audio equipments such as, radio, tape recorders etc. These are the most economical variable resistors compared to other types.

Single turn wire-wound potentiometers

The type of resistive material used in making carbon potentiometers and the method used in manufacturing, makes carbon potentiometer unsuitable for applications where,

- resistance values are to be adjusted very precisely,
- stability of the adjusted resistance values are required to be high in spite of externally and internally generated heat.

In such applications, wire-wound potentiometers are preferred to carbon potentiometers. A comparison of the carbon and wire-wound potentiometers is given below.

Property	Carbon potentiometers	Wire-wound potentiometers
Stability	Cannot be set to accurate values. Resistance value unstable. Value changes due to heat	Can be set to accurate values. Resistance values is stable. Values does not change with temperature
Power handling capacity	Generally upto 2 watts	Can handle hundred of watts.
Wear and tear	Wears out early	Has longer life
Cost	Cheap	Costly
Physical size	Small	Big to very big

Wire-wound potentiometers are available in both rotary and sliding types as shown in Chart 1. The sliding type of wire-wound variable resistors are commonly known as sliding type Rheostats. Sliding type rheostats are generally used in laboratories while carrying out experiments. High wattage rotary type rheostats are also known as VARIAC. VARIACS are used to vary the AC mains supply voltage while conducting experiments.

Sliding and rotary type of wire-wound pots are available from 10 ohms to several hundreds of ohms with power ratings from 10 watts to more than 150 watts. These are used in variable output power supplies, light dimmers etc.,

Presets and Trimpots

These are small to very small sized potentiometers. Chart 1 shows some of the types. These potentiometers are generally used when frequent adjustments are not required. These potentiometers when used in circuits, are used to trim certain value of resistance and are not generally disturbed again. Hence these are called *Presets* or *Trimpots*. These are available in a wide range of values. However, their power handling capacity will be only of the order of a few milliwatts.

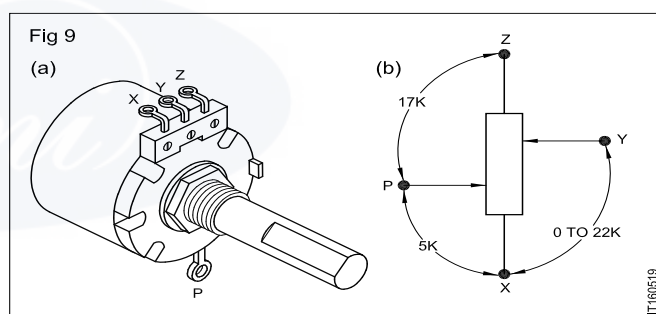
SPECIAL TYPES OF VARIABLE RESISTORS

1 Potentiometer with built in ON/OFF switch

Carbon potentiometers are also available with the shaft mechanically coupled to a ON/OFF switch. A typical potentiometer with ON/OFF switch is shown in Chart 1. The built-in ON/OFF arrangement will eliminate the need of a separate power switch in equipments/gadgets such as radios, tape recorders etc.,

2 Tapped potentiometer

Variable resistors with an additional terminal as shown in Fig 8 are available. This terminal is provided to obtain fixed resistance values in addition to variable resistance. Such variable resistors are specified for example, 5K + 17K. This example illustrated in Fig 9b.



In Fig 9b, resistance across the terminals are given below;

- X – P 5 K Ω fixed value resistance
- Z – P 17 K Ω fixed value resistance
- X – Y 0 to 22 K Ω variable resistance.

3 Ganged and Dual potentiometers

To reduce the number of controls on the front panel of the equipments, for example Volume and Tone controls of radios, television set etc., special type of potentiometers are available in the market. Two of the most popular special type of potentiometers are the *ganged potentiometer* and the *dual potentiometer*. These are shown in Chart 1.

A typical ganged potentiometer is shown in Chart 1. It consists of two potentiometer. These two pots may be of same values or different values. The control shaft of both the pots is ganged as shown in Chart. When the position of the shaft is varied, the resistance of both the pots vary.

Common applications: volume/tone control of stereos tape, players, radios and in applications where values of two resistors is to be varied simultaneously.

A typical dual potentiometer is shown in Chart 1. It consists of two potentiometers. The resistance values of these two pots may be same value or different. These two pots have independent controls but, the control shafts are built one over the other as shown in Chart 1.

Common applications : Volume/tone controls of stereos, radio, tape recorders, TVs etc., where the number of control shafts on the front panel is to be reduced.

4 MULTI-TURN POTENTIOMETERS

Most of the carbon type potentiometers (Pots) are generally of single turn type. For applications, where precise value of resistance is to be adjusted, single turn pots fail to serve the purpose. This is because the adjustment is limited to an arc of a circle only.

Hence, pots which have more than one turn and whose values can be very precisely adjusted are made. In this, the shaft of the pots can be rotated for several turns. These pots are known as *multi-turn pots*.

Multi-turn variable resistors are also available without a shaft, but with a facility to adjust the values using a small screw driver. These are called *multi-turn trimmer potentiometers* or more commonly as *multi-turn trim pots*.

Higher the number of turns the potentiometer can be rotated for a given value of resistance, higher is the precision with which the resistance value across the terminals can be adjusted.

GENERAL DEFECTS IN POTENTIOMETERS

Potentiometers become defective when used for a long period of time or when used roughly. Some of the defects that may occur in the above two cases are given in Chart 2 at the end of this lesson.

Considering the important role played by the potentiometers in any electrical or electronic circuit, it is suggested, that a defective potentiometer should be replaced instead of repairing and revising it.

While checking a variable resistor using ohmmeter, several defects can be observed. Some of these defects are listed in the table given below;

[H.I. Instructor should discuss the details of defects listed in chart 2 and table below with the trainees.]

The following care should be taken to prevent pots from becoming defective:

- Holes made for mounting trimpots and presets should be big enough such that it can be mounted easily without the other parts of the pot getting damaged.
- Do not bend the terminals of the pots to prevent terminals from getting broken.
- Do not push or force the shafts of the pot too much while fitting the knobs to prevent the wiper of the pot from getting damaged.
- Do not shake the shafts of the pots axially as the wiper of the pot may get damaged. Always use pot with knob fitted.
- Do not turn a pot beyond its maximum position as this will damage the wiper of the pot.
- Use a suitable sized screwdriver when turning presets. Oversized or undersized screwdriver may damage the slot and wiper contact with resistance track.
- Do not use high wattage soldering irons while soldering pot terminals.

POSSIBLE ELECTRICAL DEFECTS IN VARIABLE RESISTORS

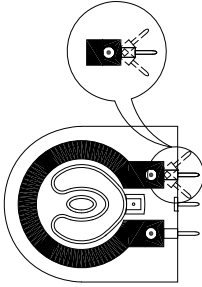
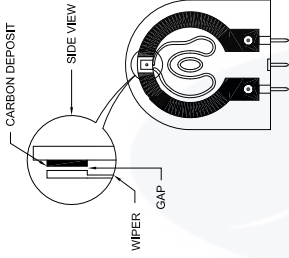
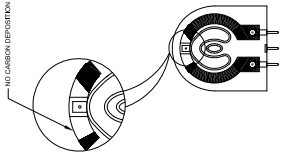
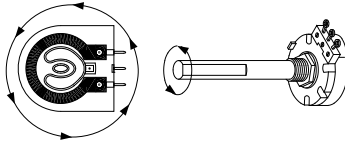
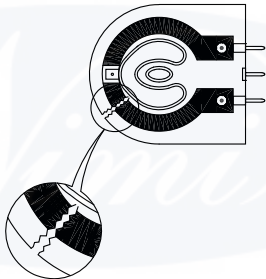
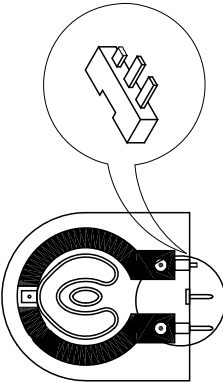
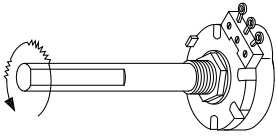
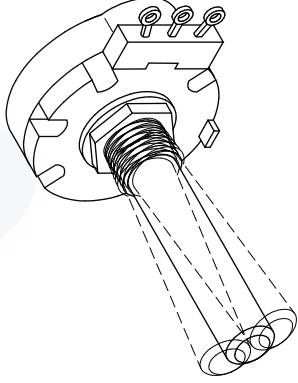
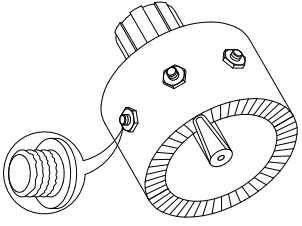
Code Number	Observation with shaft rotation	Possible defects
0	Meter pointer deflection is smooth from initial value to final value without jerks.	Nil
1	Needle deflects suddenly (jerks) from one value to an other in some places.	Wiper/Resistive Material. Material coating defective.
2	Needle suddenly goes to zero ohms or infinite. contacts defective.	Wiper/terminal
3	Needle does not move at all from zero ohm (Short) or infinite (Open) position.	Defective wiper or terminal contacts open/short.

The defect code number given in table is only for the ease of recording the defect while carrying out the exercise on variable resistors.

Chart 1

Types of Variable Resistors		WIRE WOUND	
CARBON	LONG SHAFT (Metal or plastic)	SHORT SHAFT	SCREW TYPE
	SHORT SHAFT (Metal or plastic)	TRIM POTS	SHAFT TYPE
	WITH ON / OFF	MULTI-TURN POTS	
POTS		SCREW TYPE	SHAFT TYPE
LONG SLIDE			
SHORT SLIDE			
GANGED SLIDE			
SLIDING			
RHEOSTAT			
RHEOSTAT			
GANGED			
DUAL			

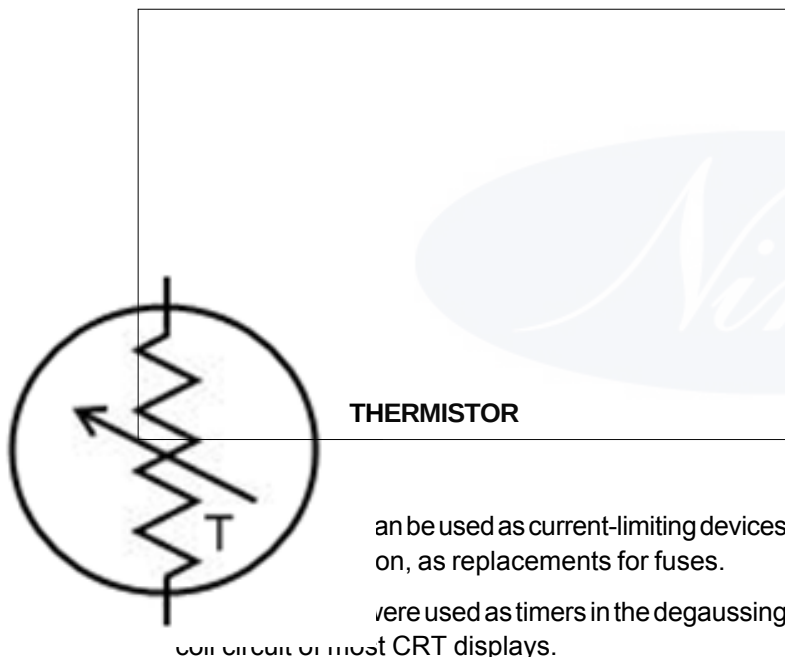
COMMON PHYSICAL DEFECTS IN VARIABLE RESISTORS

Defect code 1 TERMINAL TAG LOOSE 	Defect code 2 WIPER NOT MAKING CONTACT AT SOME PLACES. 	Defect code 3 CARBON TRACK PEELED OFF OR EXCESSIVELY WORN OUT 
Defect code 4 SLIDER DOES NOT STOP AT ENDS. POT. SHAFT ROTATES ENDLESSLY. 	Defect code 5 CRACKED BODY / BASE BOARD 	Defect code 6 TERMINALS BROKEN OR TOO SMALL TO SOLDER WIRES 
Defect code 7 SHAFT ROTATION VERY TIGHT  SHAFT SHAKEY	Defect code 8 LOOSE SHAFT 	Defect code 9 WORN OUT TERMINAL / NUT THREADS 

THERMISTOR SYMBOL

A thermistor is a resistor whose resistance changes according to the temperature it is exposed to. There are 2 main types of thermistors: negative temperature coefficient (NTC) thermistors and positive temperature coefficient (PTC) thermistors.

NTC thermistors are thermistors whose resistance decreases when the temperature it is exposed to increases. So if heat is applied to a NTC thermistor, its resistance will drop accordingly. PTC thermistors are thermistors whose resistance increases when the temperature it is exposed to increases. So if heat is applied to a PTC thermistor, its resistance increases accordingly. Being that thermistors are components which vary their resistance according to the amount of heat applied, we can exploit this property to test it. All we need to test a thermistor is an ohmmeter or a multimeter that contains an ohmmeter. We then use the ohmmeter to check the resistance after various stages of our testing to see if the thermistor is behaving according to the characteristics of the thermistor. If it is, it is good. If it does not, then it is faulty.

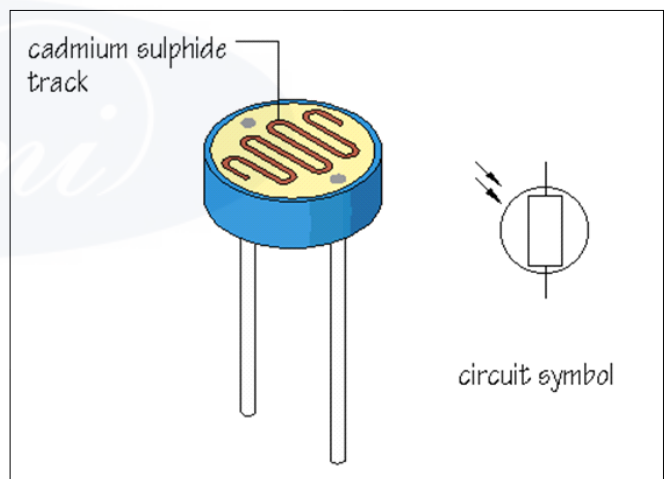


- PTC thermistors were used as heater in automotive industry to provide additional heat inside cabin with diesel engine or to heat diesel in cold climatic conditions before engine injection.
- NTC thermistors are used as resistance thermometers in low-temperature measurements of the order of 10 K.

- NTC thermistors can be used as inrush-current limiting devices in power supply circuits. They present a higher resistance initially which prevents large currents from flowing at turn-on, and then heat up and become much lower resistance to allow higher current flow during normal operation.
- NTC thermistors are regularly used in automotive applications. For example, they monitor things like coolant temperature and/or oil temperature inside the engine etc.
- NTC thermistors can be also used to monitor the temperature of an incubator.
- Thermistors are also commonly used in modern digital thermostats and to monitor the temperature of battery packs while charging.
- Thermistors are also used in the hot ends of 3D printers, they produce heat and keep a constant temperature for melting the plastic filament.

LIGHT DEPENDENT RESISTOR(LDR)

A light dependent resistor (LDR) or photoresistor is a resistor whose resistance decreases with increasing incident light intensity; in other words, it exhibits photoconductivity.



A photoresistor is made of a high resistance semiconductor. If light falling on the device is of high enough frequency, photons absorbed by the semiconductor give bound electrons enough energy to jump into the conduction band. The resulting free electron (and its hole partner) conduct electricity, thereby lowering resistance. There are many types of photoresistors, with different specifications and models. Photoresistors can be coated with or packaged in different materials that vary the resistance, depending on the use for each LDR.

Applications

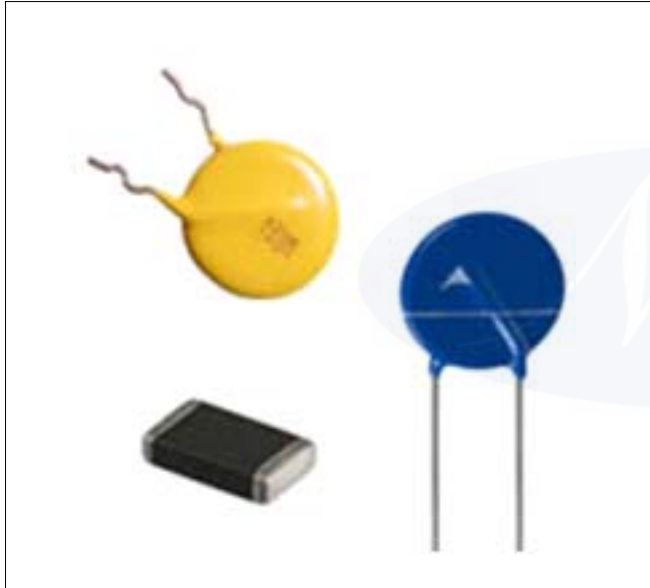
Photoresistors come in many types. Inexpensive cadmium sulphide cells can be found in many consumer items such as camera light meters, street lights, clock radios, alarm devices, outdoor clocks, solar street lamps and solar road studs, etc.

They are also used in some dynamic compressors together with a small incandescent lamp or light emitting diode to control gain reduction and are also used in bed lamps, etc.

VOLTAGE DEPENDENT RESISTOR / Varistor (VDR)

VARISTOR SYMBOL

The resistance of a varistor is variable and depends on the voltage applied. The word is composed of parts of the words "variable resistor". Their resistance decreases when the voltage increases. In case of excessive voltage increases, their resistance drops dramatically. This behavior makes them suitable to protect circuits during voltage surges. Causes of a surge can include lightning strikes and electrostatic discharges. The most common type of VDR is the metal oxide varistor or MOV.



VARISTOR SYMBOL

Applications

- Telephone and other communication line protection
- Radio communication equipment transient suppression
- Surge protector power strips
- Cable TV system surge protectors
- Power supply protection

Working principle of oscilloscope

Objectives : At the end of this lesson you shall be able to

- state the meaning of AC current and AC voltage
- state the method of graphically representing an electrical quantity
- describe the important terms used in quantifying AC
- state the names and functions of front panel control switches and inputs/outputs of a typical CRO
- state the method of measuring waveform using a CRO.

Alternating current

Alternating current is a current that changes the direction of flow periodically, changing its instantaneous magnitude.

If an alternating current is connected to a load, then an alternating voltage is developed across the load. We often refer to this developed alternating voltage as *AC voltage*. The other way round, an AC voltage produces an alternating current through a load. Likewise, AC power refers to the power that is produced or consumed by alternating current and alternating voltage.

Waveforms

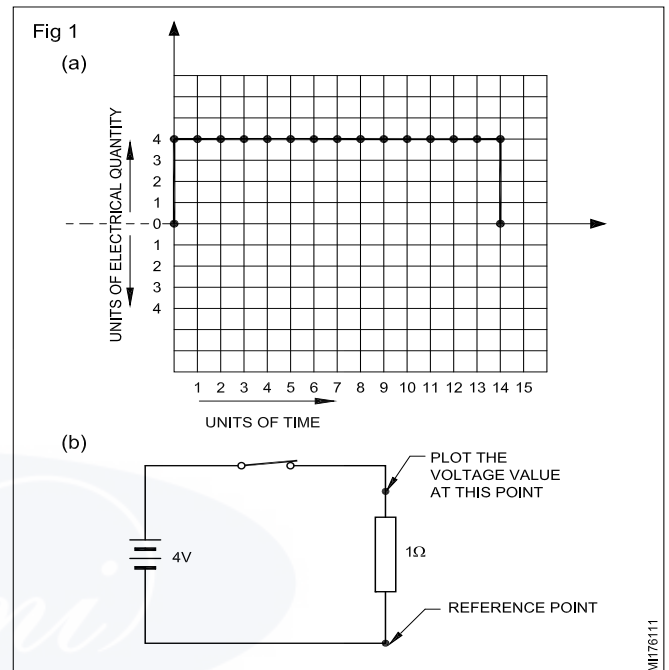
An electrical waveform could be represented by a line or a curve on a graph, as shown in Fig 1a. The line or curve is produced by plotting points on a graph and then connecting the points. The points represent the value of some electrical quantity at different times. The magnitude and direction of the electrical quantity (voltage or current) are generally indicated on the vertical axis of the graph. Time is indicated on the horizontal axis.

The waveform shown in Fig 1a represents the voltage across the resistor in Fig 1b. In this case, the units on the vertical axis is volts. The waveform shows that when the circuit is energised, the voltage rapidly increases to its maximum value (4V). The voltage remains at its maximum value until the circuit is opened. When the circuit is opened at the 14th unit of time, the voltage across the resistor suddenly drops to zero.

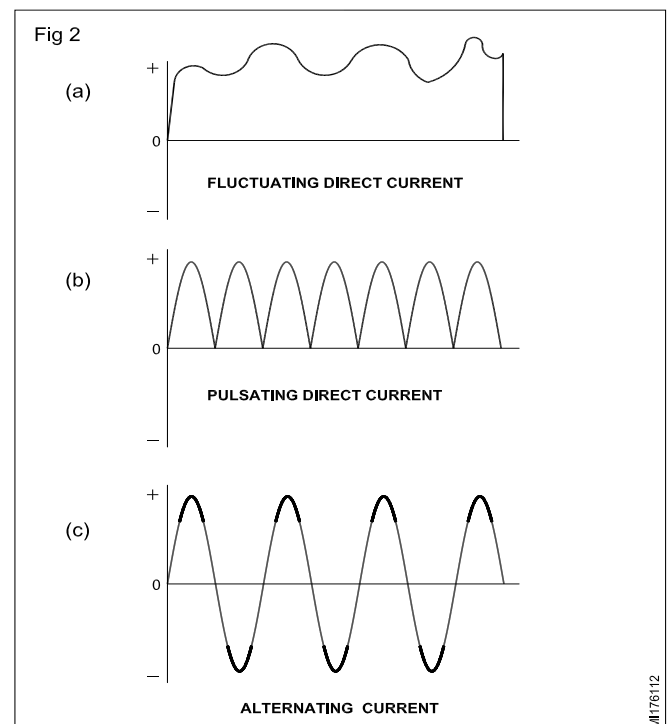
The waveform of Fig 1a could just as well represent the current in Fig 1b. For the current waveform the unit for vertical axis will be amperes.

Waveform polarity

In Fig 1a as can be seen, the waveform is always above the zero reference line. This means that the polarity of the plotted voltage in Fig 1b is always positive. If the waveform is in case a current waveform, then the waveform means, that the direction of the current is unidirectional or it never reverses. In other words, Fig 1a is a waveform of a steady DC voltage or DC current.



Waveforms other than pure DC waveforms are shown in Fig 2a, 2b and 2c. The most common type of AC waveform is shown in Fig 2c. This waveform is referred to as sinusoidal or sine wave.

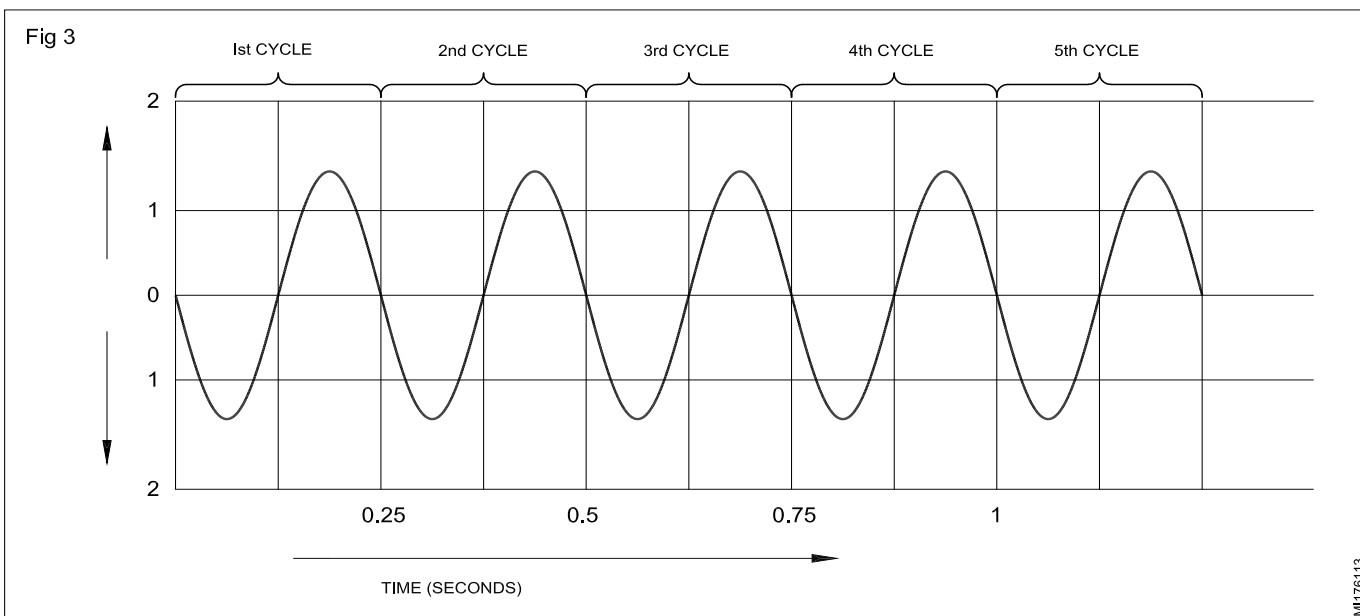


Quantifying AC

Full description of an alternating sinusoidal voltage or current requires the use of a number of terms. Some of these terms have general meanings, but, when used in electricity they have very specific meanings. The most important terms used with AC voltage and AC current are given below;

CYCLE

The waveform in Fig 3 shows five cycles of AC. Each cycle in Fig 3 is a duplicate of every other cycle in the figure.



The part of the cycle above the horizontal line or the zero reference line in Fig 3 is called **positive half cycle**. A half cycle is also called an **alternation**. Therefore, the positive half cycle could be called the positive alternation. The **negative half cycle**, is that part below the horizontal reference line.

PERIOD

In any AC waveform, the time required to complete one cycle is called the **period (T)** of a waveform. In Fig 3 the period (T) of that waveform is 0.25seconds.

FREQUENCY

The number of complete cycles produced in one second is called the **frequency (f)** of an AC current or AC voltage. Therefore, frequency refers to how rapidly the current or voltage changes its polarity. In Fig 3, the frequency (f) of that waveform is 4 cycles or 4 Hz.

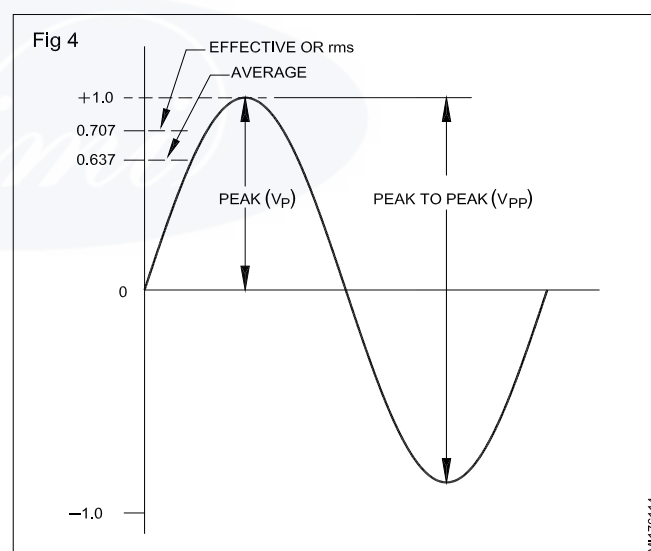
Specifying amplitude of AC waveforms

The amplitude of an AC waveform can be specified in several different ways, as shown in Fig 4. These ways of specifying amplitude are common for both voltage and current.

PEAK and PEAK-TO-PEAK VALUE

For symmetrical waveforms like sinusoidal waveform, the **peak value (V_p)**, also called the maximum value, of both the positive and negative half cycle will be equal. Therefore, the **peak-to-peak value (V_{p-p})** is twice as great as the peak value. For a sinusoidal AC voltage we can write,

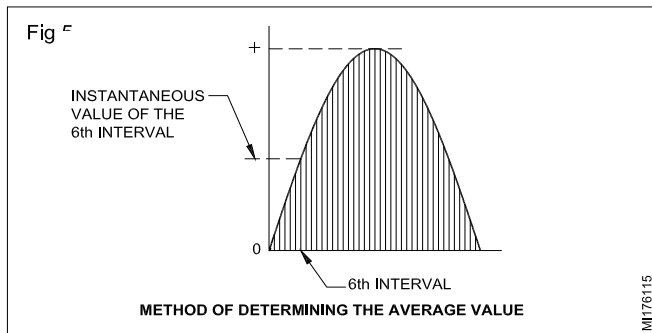
$$V_{p-p} = 2 V_p \quad \dots\dots[1]$$



Average value

The **average value (V_{av})** of a waveform is the arithmetical mean value. Finding the arithmetical mean of a +ve half cycle of a sine waveform is illustrated in Fig 5.

In Fig 5, the instantaneous values of the +ve half cycle of a sinewave waveform is determined at a number of equally spaced intervals along the horizontal axis. Then, all the instantaneous values are added together and divided by the total number of intervals. The closer the intervals, the more accurate the estimate of the average value of the waveform. In a pure sinusoidal waveform, since the +ve and -ve half cycles are the same the average value of the



The relationship between peak value and average value is indicated in Fig 4. This relationship can be written as,

$$V_{av} = 0.637 V_p \quad \dots [2]$$

$$\text{or, } V_p = 1.57 V_{av} \quad \dots [3]$$

The relationship between average and peak is used when analyzing circuits which convert AC voltage to pulsating DC voltage.

Example: What is the average voltage of the +ve half cycle of a sinusoidal waveform if its peak-to-peak voltage is 300V?

Given: $V_{p-p} = 300V$

To find: V_{av}

known: $V_{av} = 0.637 V_p$

Solution:
$$V_p = \frac{V_{p-p}}{2}$$

$$V_p = \frac{300 V}{2} = 150 V$$

$$V_{av} = 0.637 \times 150 V = 95.6 V$$

Answer: The average voltage of the +ve half cycle is +95.6V.

The average value of the negative half cycle of the same waveform will be –95.6V. Hence the average value of one complete cycle of this sinewave is equal to 0.

Root-mean-square (rms) or effective value

Another most common way of quantifying AC is by stating its effective, or rms (root mean square) value. The rms value of an alternating current, is given by that steady DC current, which when flowing through a given circuit for a given time produces the same heat as produced by the alternating current by flowing through the same circuit for the same time.

For sinusoidal AC, the rms value and peak value are related by the following formulae;

$$V_{rms} = 0.707 V_p \quad \dots [4]$$

$$\text{and } V_p = 1.414 V_{rms} \quad \dots [5]$$

Example: What is the peak voltage of 230 V rms main AC voltage?

Given: $V_{rms} = 230 V$

To find: V_p

Known: $V_p = 1.414 V_{rms}$

Solution: $V_p = 1.414 \times 230 V = 325.22 \text{ volts}$

Answer: The peak voltage of main AC is 325.22 V.

AC currents and AC voltages are usually specified in rms values. It is a common practice to assume that AC is in rms units unless a subscript “p-p” or “p” is included. For example, the voltage at the outlets in our home is specified as 230 V. This means, it is 230V rms voltage.

Waveform measuring instrument - THE CATHODE RAY OSCILLOSCOPE (CRO)

The cathode ray oscilloscope abbreviated as CRO is the most common instrument for the study of waveforms. It can be used for observing the type of waveform and for measuring voltage, frequency, time interval, and phase etc.

Chart 1 at the end of this lesson shows the front panel and controls of a simple, low cost oscilloscope.

Make a chart of the front panel and controls of the oscilloscope available in your institute for ease of understanding and use during exercises.

These oscilloscopes are generally powered by 230 Volts, 50 Hz, AC mains supply. Important parts of a CRO and their functions are given below;

SCREEN

The waveforms under study are displayed on the screen. This screen on its inside surface is coated with fluorescent chemicals that glow when struck by a beam of electrons. The screen is protected by a flat piece of transparent hard plastic, called **graticule** which has vertical and horizontal calibrated lines like a graph sheet.

ON/OFF and INTENSITY CONTROL

The power ON/OFF switch for low cost CRO's are generally incorporated along with the INTENSITY control. The INTENSITY control allow the brightness of display to be adjusted.

FOCUS CONTROL

The FOCUS controls permit the beam to be focused to a fine point or to a fine line when a waveform is being displayed.

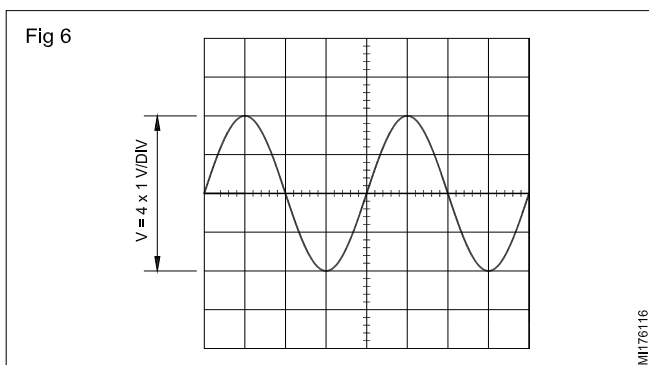
VERTICAL(Y) BEAM POSITION CONTROL

The Y-POSITION controls are used to move the displayed waveform or beam vertically up or down the screen, just to set it in a convenient position for viewing.

AMPLITUDE/DIVISION OR SENSITIVITY SELECTOR

The AMPL/DIV selector switches determine the sensitivity of the display to input voltages. When this control is set to 1V, that is, 1V per vertical division on the graticule, an input voltage with a peak-to-peak amplitude of 4 V, for example, would occupy four vertical divisions on the oscilloscope screen as shown in Fig 6.

In some oscilloscopes a small control knob at the centre of the AMPL/DIV control knob will be provided. This knob will permit continuous sensitivity adjustment, so that the display amplitude may be increased as desired. When the AMPL/DIV selection is to be used for measuring the



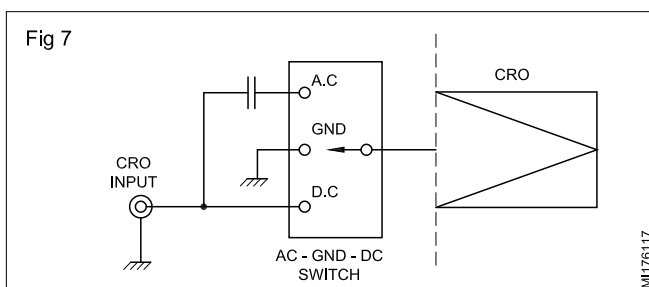
amplitude of a displayed waveform, the small centre knob must be turned to its extreme clockwise or anticlockwise (calibrated) position depending on the oscilloscope.

This knob is used to calibrate the AMPL/DIV when the calibration gets offset.

The typical sensitivity range available (as illustrated in chart 1) is from 2mV/division to 10V/division. In some oscilloscopes a “1 x 10” switch alongside the AMPL/DIV selector knob will be provided which allows the sensitivity to be multiplied 10 times. This selector knob is used when a higher amplitude waveform is to be displayed. Thus, the amplitude of the displayed waveform can be increased 10 times, and this tenfold increase applies to all sensitivity selections.

AC, DC, O(GND) switches

As shown in Fig 7, the AC, DC and O(ground) switches permit the input signal to be connected to the CRO input **directly (DC position)**, **CRO input grounded (O position)**, or connected **via a capacitor (AC position)**.



Capacitors are electronic components which allows AC to pass through it and blocks DC.

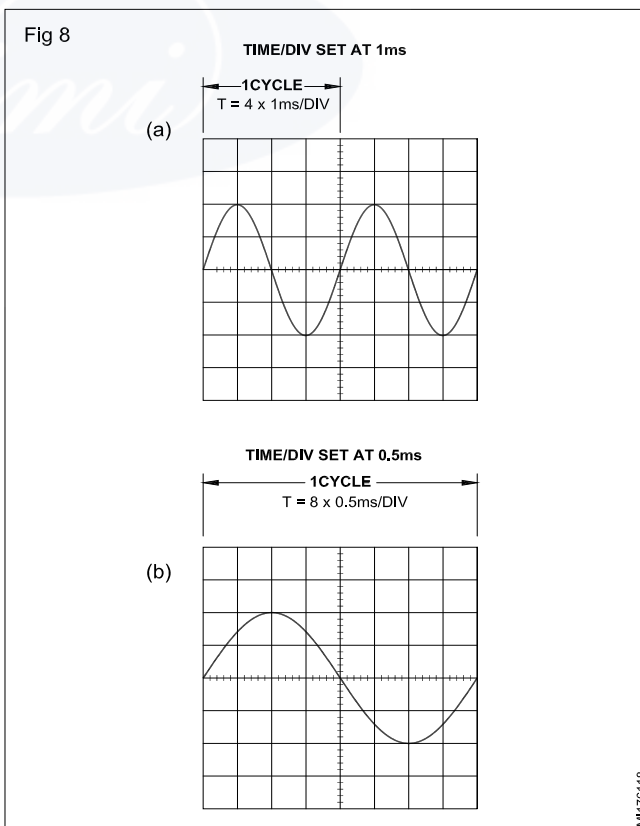
Direct connection would be required, for example, to study DC voltage levels, or to investigate AC-DC voltage combinations. Sometimes it is necessary to display an AC voltage but not the DC level it may be superimposed upon. This is achieved by setting the switch to AC, where the coupling capacitor passes the AC quantity and blocks DC.

SIGNAL INPUT PORTS

The coaxial-type input connections (BNC-sockets) for each channel is given on the front panel as A and B. Two separate signals can be given simultaneously to these two inputs. This is specially useful while comparing two waveforms and while measuring phase shift.

TIME/DIVISION or TIME BASE

As already explained, the AMPL/DIV switch adjusts the vertical amplitude of the display, by selecting the voltage sensitivity. In a similar way, the TIME/DIV switch or *time base* determines the horizontal amplitude of each cycle of input waveform displayed. This control applies to both channels (A & B). Refer to the waveform displayed on the screen in Fig 8a and assume that the TIME/DIV switch is set to 1 ms (i.e., 1 ms/horizontal division on the graticule). One cycle of the waveform occupies four horizontal divisions; therefore, its *time period* is $4 \times 1 \text{ ms}$, or 4 ms. If the TIME/DIV switch is set to 0.5 ms (i.e., 0.5 ms/horizontal division on the graticule) as shown in Fig 8b, one cycle of the wave form occupies eight horizontal divisions.



A typical time-base range available as shown in chart 1, is from 0.5 $\mu\text{s}/\text{division}$ to 200 ms/division. At one extreme position of the TIME/DIV knob, TRIG or X EXT marking will be made. When put to this position, the internal circuits which produce horizontal motion of the electron beam will get disconnected. This allows horizontal deflection to be controlled by an external voltage connected to the CRO at the terminal marked as TRIG or X-EXT.

In some oscilloscopes a small control knob at the centre of the TIME/DIV knob will be provided for continuous time-base adjustment, so that the displayed waveform may be increased horizontally (widened) as desired. When the TIME/DIV selection is to be used for measuring the time period of a displayed waveform, the small centre knob must be turned to its extreme clockwise or anticlockwise (calibrated) position depending upon the oscilloscope.

This knob is used to calibrate the TIME/DIV when calibration gets off-set.

TRIGGERING MODE - LINE(A/B), EXT, $\pm$, NORMAL/TV

In order to provide a useful display of waveforms, it is necessary to ensure that the display commences at the left hand side of the screen, exactly when one of the input waveforms is at its zero position. Thus, the time-base circuits are said to be triggered at this instant.

A - LINE - B positions

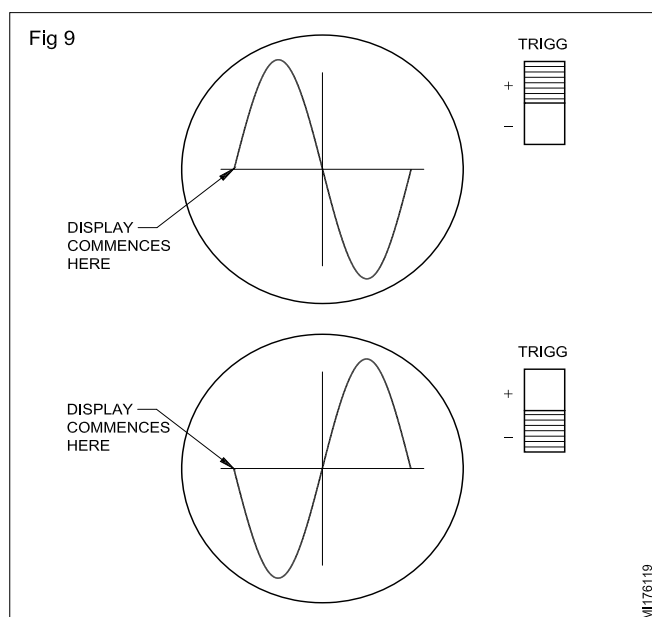
The triggering waveform is selected by the A - LINE - B switch. Thus, at A, the waveform on beam A (or channel A) is used to trigger the time base. Whereas at B position, the waveform on beam B triggers the time base.

EXT position

When put to EXT, triggering is provided by an external source connected at the TRIG input terminal.

$\pm$ position

Triggering may be + or - , selected by the $\pm$ switch. This simply means that the displayed waveform may be made to commence either when it is *going positive* or when it is going negative as shown in Fig 9.



NORMAL/TV position

The NORMAL/TV switch applies only for certain tests on television sets and for most purposes is left at the NORMAL position.

TRIG LEVEL CONTROL

The LEVEL knob is a triggering level control knob for the time-base circuits. This determines the voltage level of the trigger waveform at which the display commences. By adjusting the knob, an observable stationary waveform can be obtained on the screen.

HORIZONTAL(X) POSITION CONTROL

The X position control has a similar function to the Y position knobs. It shifts the displayed waveforms horizontally about the screen, as desired for the best viewing.

Some oscilloscopes have the X MAGN (X-magnification) knob which varies the sensitivity of the TIME/DIV selection. When the X MAGN is at extremely counterclockwise position (at x 1), all the TIME/DIV selections are exactly as identified. When the X MAGN control is at its extreme clockwise position (x 5), the TIME/DIV selections all have their sensitivities increased five times. X MAGN control is usually kept at its x 1 position while making measurement.

When making measurement on electronic circuits, it is often a great advantage to be able to display two different waveforms simultaneously on the CRO screen to find the relationship between these waveforms.

To enable viewing two waveforms simultaneously a control switch generally called ALT/CHOP (Alternate/chopped) is available on the front panel of CRO's.

In alternate(ALT) mode the two waveforms are displayed alternately on the CRO screen. However, due to persistence of vision, both the waveforms are seen simultaneously on the screen.

In chopped (CHOP) instead of alternative display of the full waveform only small portions of both the waveforms are displayed consecutively and alternatively. In CHOP mode the waveforms will look much brighter than in ALT mode.

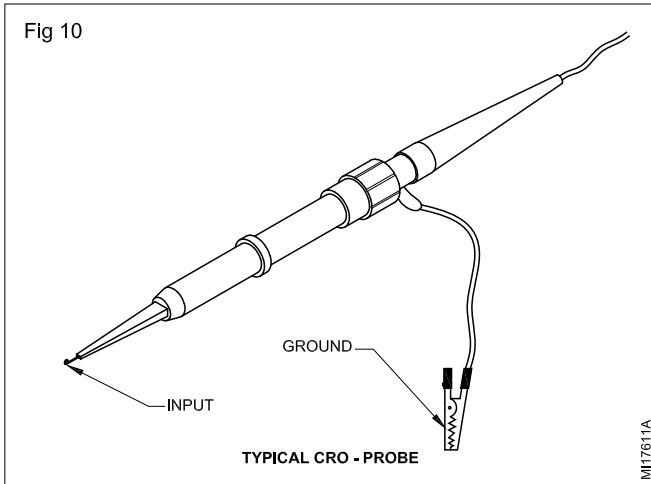
GRATICULE ILLUMINATION CONTROL

Some oscilloscopes have ILLUM switch which illuminates the graticule to enable easy reading of the waveform levels. By this the level of illumination of the graticule can also be varied depending on requirement.

OSCILLOSCOPE PROBES

The signals to be displayed on CRO are normally connected to the oscilloscope via coaxial cables called CRO probes. One end of a CRO probe will have a BNC female connector and the other end will have an insulated connecting clip as shown in Fig 10.

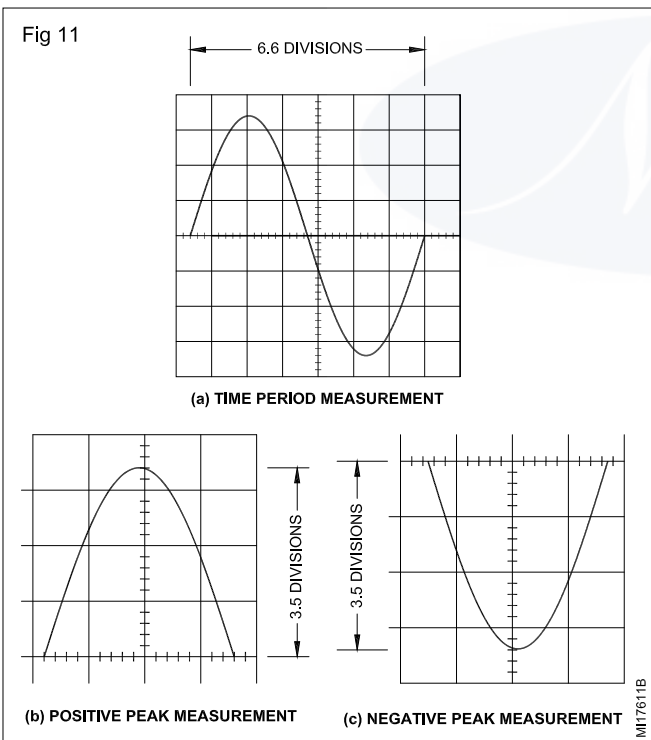
Fig 10



Some probes termed as attenuator probes, have resistors inside them to increase the input impedance from the normal 1-MW input impedance of the oscilloscope to 10 MW. These probes also have the effect of reducing the voltage applied to the oscilloscope by a factor of 10. Thus, they are usually referred to as 10:1 probes, and the ordinary non-attenuator probes are termed 1:1 probes.

Fig 11a illustrates the method that should be used to make accurate measurement of displayed signal parameters.

Fig 11



Steps to follow;

- By means of the appropriate Y SHIFT control, the unused beam should be shifted off the screen.
- The AMPL/Div and TIME/Div controls are then to be adjusted to give the largest possible display of one cycle of input waveform as shown in Fig 11a. Note that the small central knobs on the AMPL/div and TIME/div controls if available must be turned to calibrated position.

From Fig 11a, the time period of the waveform is 6.6 divisions x time/division setting.

Fig 11b and 11c shows how the displayed waveform should be moved horizontally by the X position control, in order to accurately measure the peak or peak-to-peak amplitude of the waveform.

From the Fig 11b and 11c, the peak-to-peak is

3.5 + 3.5 divisions x volts/division setting.

For specific details about the front-panel controls and inputs/outputs of the CRO used in your institute, read the CRO user manual by obtaining the same from your instructor.

AUDIO SIGNAL GENERATOR

Audio signal generators provide sinusoidal output voltage. The frequency of the output signal is usually variable from 1Hz to several tens of kilo Hz. The amplitude of output is also variable with a maximum value ranging from 10 to 20 volts peak-to-peak. Signal generators are used to inject signals into circuits such as amplifiers, allowing oscilloscope to make measurements of the gain and frequency characteristics of the circuit. Audio signal generators are generally referred to as AF generators. A typical audio signal generator is shown in Fig 12.

Fig 12

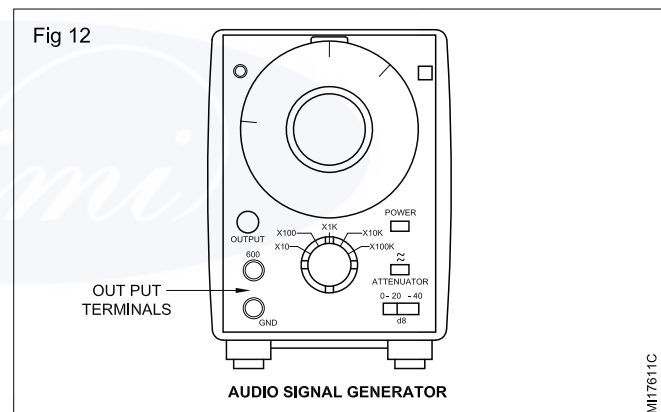
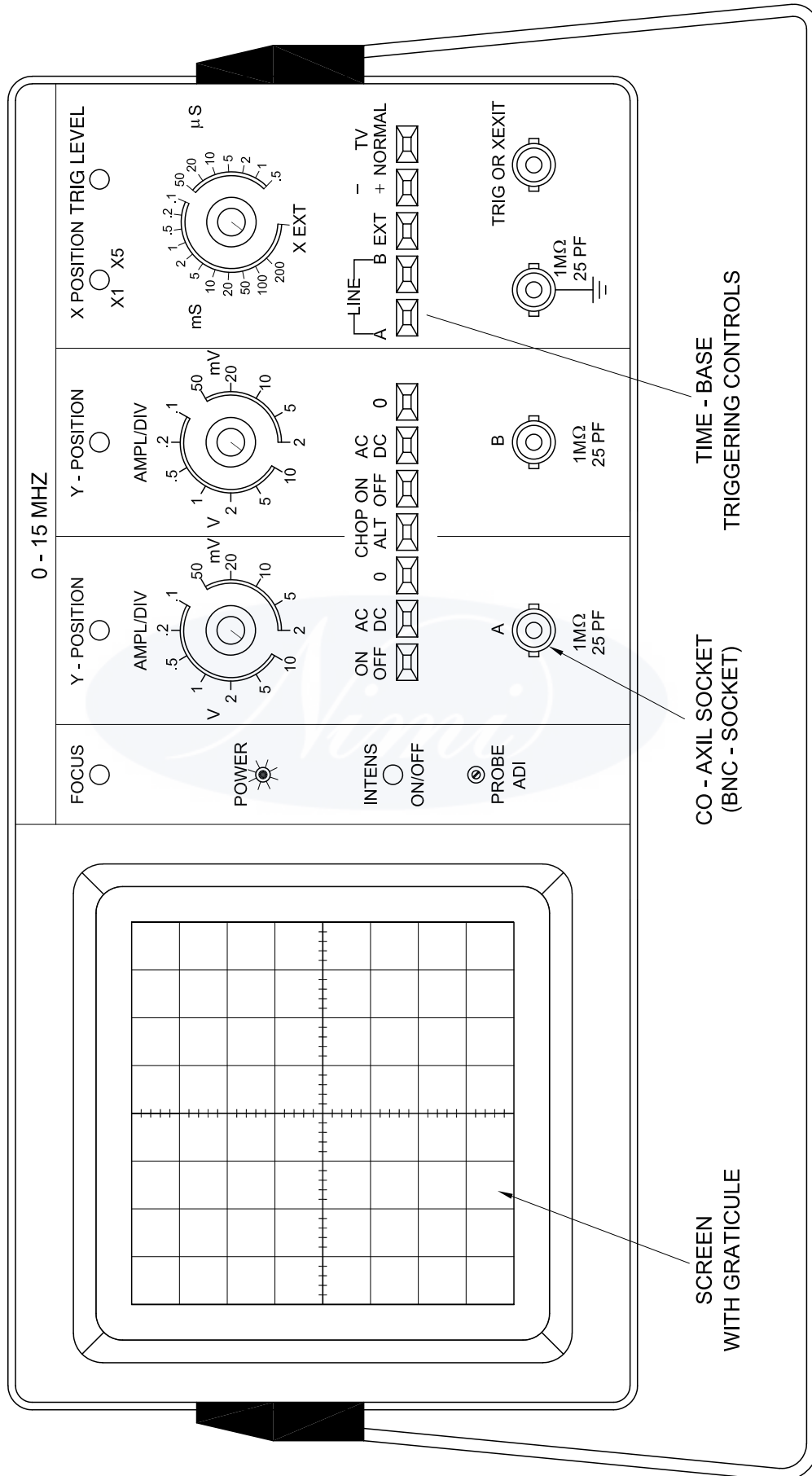


CHART - 1

Typical Oscilloscope (CRO) and Controls, Inputs and Outputs

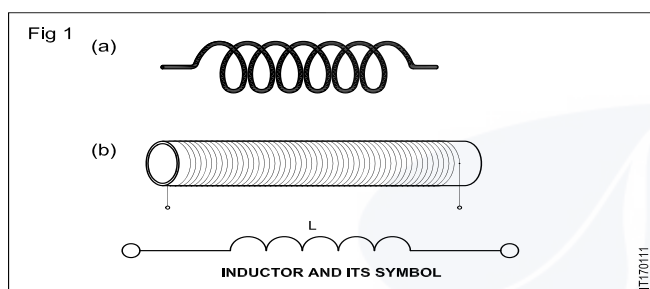


Inductors and Inductive Reactance at different frequencies

Objectives: At the end of this lesson you shall be able to

- state meaning of inductor and inductance
- explain through simple example how inductor opposes sudden changes in current through it
- state the meaning of self induction and unit of inductance, its definition
- list factors determining the value of an inductor and formula for finding inductance of a coil
- explain the meaning of *inductive kick*
- list the different classification of inductors and list few types, applications
- calculate the resistance of a given wire length and meaning of X_L
- explain the phase relationship between current & voltage in a purely inductive circuit
- list the factors affecting inductive reactance and power consumed by pure inductance
- state the meaning of impedance of a R-L circuit and explain the power triangle
- state the meaning and importance of Q of coil.

Inductors are components consisting of coil of wire as shown in Fig 1. The coil of wire may or may not be wound on a former as shown in Fig 1a and 1b. The basic function of an inductor is to **store electric energy in the form of magnetic field**, when current flows through the inductor.



Inductance is the electrical property of inductors. Letter 'L' is used as a symbol to represent Inductance. Inductance, is the ability of a device to oppose any change in the current flowing through it. This opposition to change in current, is achieved by the energy stored by it, in the form of magnetic field.

Inductance, and thus an inductor, chokes off or restricts sudden changes in current through it. The change may be either increasing or decreasing. Hence inductors are also sometimes called as Chokes.

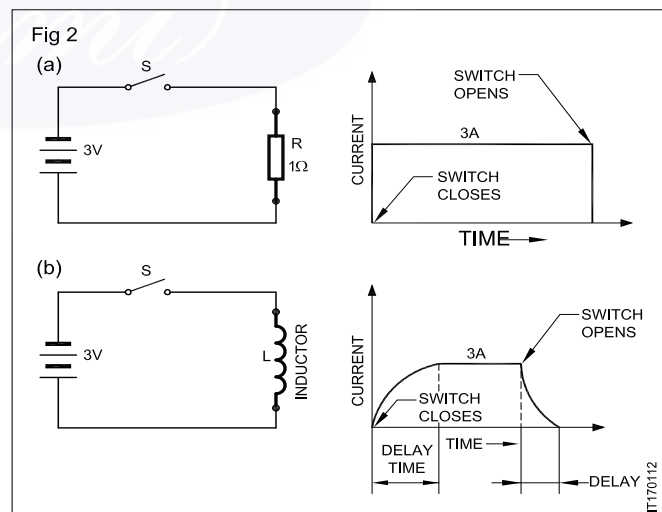
Principle of operation

Recall that, when current begins to flow through a conductor, magnetic flux rings start to expand around the conductor. This expanding flux induces a small voltage in the conductor called back-emf or counter emf. This induced voltage has a polarity that opposes the source voltage which creates the induced voltage.

Thus, the inductance in a coil of wire, carrying current, opposes any rise or fall of current through it and tries to keep the current through it constant.

It should be noted that, the inductance cannot completely stop the increase in current because, the induced voltage is caused by the increasing flux, and the increasing flux depends on the increasing current. Therefore, an inductor can restrict only, the rate at which the current can increase or decrease through it.

Example: A Resistor of 1Ω is connected to a DC source of 3 volts, as shown in Fig 2a. The moment switch S is ON, current will increase from 0 to its steady state value of 3Amps instantaneously, as shown in graph. When the switch is opened, the current drops back to

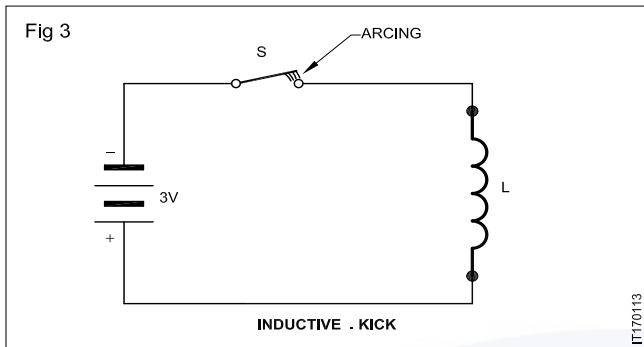


zero just as fast as it raised.

Whereas, when the same DC voltage is applied to an Inductor having a coil resistance of 1Ω as shown in Fig 2b, the current will not increase instantaneously from 0 to its steady value because the inductor in the circuit does not allow it to happen. The current will reach the steady state value after a time delay as shown in graph. The amount of delay depends on the value of inductance and the ohmic resistance of the inductor.

Once the current through the circuit in Fig 2b reaches its steady state value of 3Amps, which is decided by the ohmic value of the inductance, the magnitude of current remains constant and hence the inductive effect stops. At this point, the only opposition the inductor offers is its ohmic/DC resistance.

When the switch S of Fig 2b is opened, the back-emf(bemf) or counter emf(cemf) of the inductor becomes very high, much greater than the source voltage. This high voltage(cemf), prevents the current from instantaneously dropping to zero. It does this by ionizing the air between the switch contacts as the switch opens. This causes the switch contacts to arc and burn as shown in Fig 3. This is known as *inductive kick*. As the energy stored in the inductor's magnetic field gets used up, the switch contacts deionize and current stops.



This property of a coil to induce an emf within the coil due to a changing current through it is termed as **SELF INDUCTANCE**.

Unit of inductance - The Henry

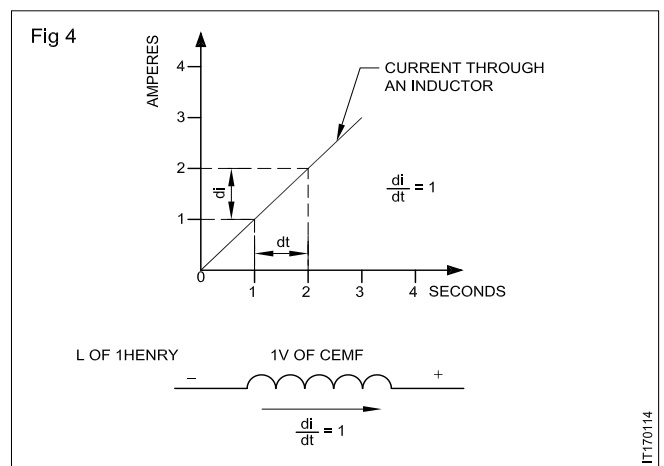
The basic unit of measure of Inductance is **Henry** abbreviated as **H**. The unit henry is defined in terms of, the amount of cemf produced when the amplitude of current through the inductor is changing. Based on this, One Henry is that amount of Inductance which develops 1 V of cemf in the coil when the current changes at the rate of 1 Amp/sec.

From the above definition, referring Fig 4,

$$\text{Inductance, } L = \frac{V_L}{di / dt}$$

Where, V_L = Induced voltage

and $\frac{di}{dt}$ = rate of change of current. Refer Fig 4.

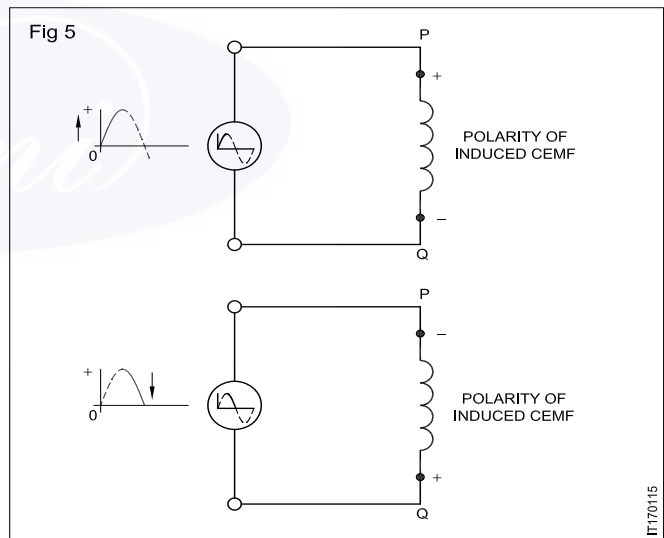


Polarity of Induced emf

The induced emf (voltage) in an inductor (cemf) has polarity that always opposes the source voltage (Lenz's law).

Fig 5 shows an inductor across an ac voltage source. When the applied voltage is increasing from 0 to +ve peak as shown in Fig 5a, the counter emf at end P of inductor will have +ve polarity opposing the increasing source voltage.

In Fig 5b, when the source voltage is decreasing from +ve peak to zero, the cemf at end P of the inductor will have -ve polarity opposing the decreasing source voltage.

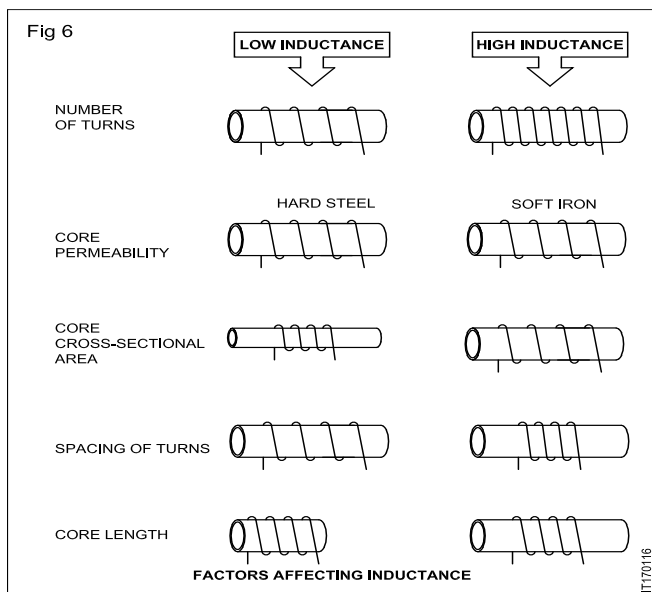


Factors determining value of Inductance

The inductance of an inductor is primarily determined by the following four factors:

- 1) The number of turns of wire
- 2) The material on which the coil is wound or the core material
- 3) The spacing between turns of wire and
- 4) The diameter of the coil

Fig 6 illustrates the effect of these factors on the inductance value.



Given the parameters listed above, the inductance of a coil can be calculated using the formula,

$$L = \mu \frac{N^2 A}{l} \text{ Henries}$$

where,

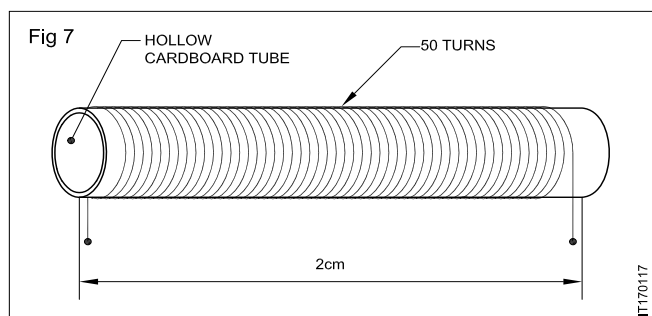
μ = Permeability of the magnetic core around which the coil is wound, in Wb/At-m ($\mu = \mu_0 \mu_r$)

N = Number of turns of the coil

A = Area of cross-section of the core in square metres, m^2

l = length of the coil in meters.

Example: In Fig 7, a coil of 50 turns is wound on a hollow card board tube of cross sectional area 1 cm^2 . The tube is 2 cms long. Calculate the inductance of the coil.



SOLUTION:

Since the card board tube is non magnetic and hollow, its permeability (μ) is nothing but permeability of air.

Therefore in this case $\mu = \mu_0 = 4\pi \times 10^{-7} \text{ Wb/At-m}$

Area of cross section = $1 \text{ cm}^2 = 0.0001 \text{ m}^2 = 1 \times 10^{-4} \text{ m}^2$

Length of the core = $2 \text{ cm} = 2 \times 10^{-2} \text{ m}$

From above data,

$$\text{Inductance of the coil, } L = \mu \frac{N^2 A}{l}$$

$$= \frac{4\pi \times 10^{-7} \times 50^2 \times (1 \times 10^{-4}) \text{ m}^2}{(2 \times 10^{-2}) \text{ m}} = 15.7 \times 10^{-6} \text{ H}$$

Values of inductance less than 1 Henry can be represented as follows;

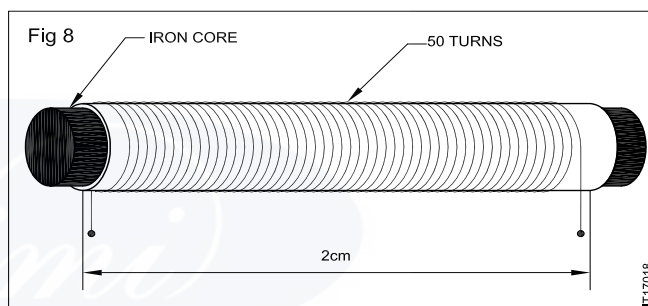
$$1 \times 10^{-3} \text{ Henries} = 1 \text{ milli Henry} = 1 \text{ mH}$$

$$1 \times 10^{-6} \text{ Henries} = 1 \text{ micro Henry} = 1 \mu\text{H}$$

$$1 \times 10^{-9} \text{ Henries} = 1 \text{ nano Henry} = 1 \text{ nH}$$

Hence the value of inductance calculated in the example above was $15.7 \mu\text{H}$.

In the above example if an iron core of relative permeability (μ_r) 200 is inserted in the hollow core as shown in Fig 8, calculate the new value of inductance.



$$\text{The changed permeability of core is } \mu = \mu_r \times \mu_0 = 200 \mu_0.$$

Hence, the new value of L is,

$$= \frac{200 \times 4\pi \times 10^{-7} \times 50^2 \times (1 \times 10^{-4}) \text{ m}^2}{(2 \times 10^{-2}) \text{ m}} = 3.1 \text{ mH}$$

A mere insertion of the iron core increases the inductance by a factor of almost 200. This indicates the major influence of core permeability on inductance value of an inductor. Higher the permeability of the core material used, higher will be the inductance value of an inductor.

In the above example, if the number of turns of coil is doubled, i.e., increased from 50 to 100 turns, inductance value will increase four fold. The new inductance value would be $L = 4 \times 3.1 \text{ mH} = 12.4 \text{ mH}$.

The permeability of materials, iron for example varies widely with the amount of flux density (B) and naturally the current in the coil. This means that the inductance of any coil with a magnetic core is never constant. The inductance value varies with the amount of current (flux) in the coil. For this reason many coil manufacturers quote the value of inductance at some specified current. For example, a coil may have 8H of inductance at 85 mA, but if operated at only 1 mA, the inductance would be 14H.

PRACTICAL INDUCTORS AND TYPES

For practical applications, inductors are manufactured to give a specified amount of inductance. Value of practical inductors range from a few microhenries for application in high frequency communication circuits upto several henries for power supply ripple filter circuits.

Inductors can be classified under various categories as shown in Chart-1 given at the end of this lesson. Illustration of different types of inductors are shown in Chart-2 at the end of this lesson.

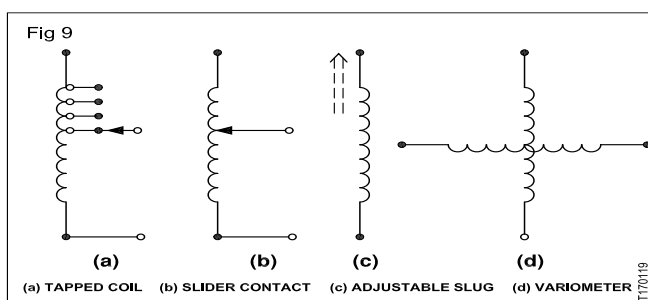
Air core coils have practically no losses from eddy currents or hysteresis. However inductor with air core have their values limited to low values in the range of micro to milli Henries. Air core inductors are used in high frequency applications.

Laminated Iron Core is formed using a group of individual laminations. Each lamination is insulated by a thin coating of iron oxide, silicon steel or varnish. This insulation increases the resistance reducing eddy current losses. These type of inductors are generally used for mains frequency of 50/60 Hz and lower audio frequency range, upto 10 KHz.

Powdered Iron Core is used to reduce the eddy currents in the core when used at radio frequencies. It consists of individual insulated granules pressed into one solid form called slug.

Ferrite Core is made from synthetic ceramic material which are ferromagnetic. They provide high value of flux density like iron, but have the advantage of being insulators, thus reducing the eddy current losses to bare minimum. Because of this advantage, inductors with ferrite core are used for high to very high frequency application.

Variable Inductors unlike fixed Inductors, variable inductors have the facility to vary its inductance value either in steps or continuously. The inductance can be varied by any one of the methods as shown in Fig 9.



Methods (a) & (b) shown in Fig 9 are generally used in low frequency application.

Method (c) is used in coils with ferrite as core. These are used in high frequency applications.

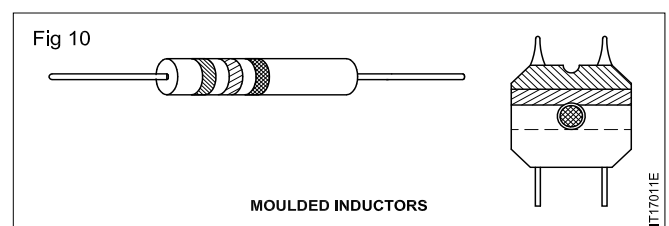
Method (d) is known as Variometer arrangement. In this, the position of one coil is varied within the other. The total inductance will be minimum when the coils are perpendicular to each other.

Shielded/Screened inductors will have a metal cover over the inductor. The shield is usually made of copper or aluminum. The reason for shielding is to isolate the coil from external varying magnetic field and to minimize the effect of the coils RF current on external circuits.

While making a shield/screen for an inductor the following points are to be noted;

- metal used as cover should be a good conductor
- clearance between the sides of the coil and the metal should be equal to or greater than the coil radius. If the clearance is less, the shield reduces the inductance value drastically.

Moulded inductors, looks like resistors with their values colour coded. The coding scheme is same as in resistor, except that the value of L are given in microhenry (μH). For example, a coil with yellow, red and black stripes or dots as shown in Fig 10, has inductance value of 42 μH .



Laboratory type variable inductor are available in the form of a **decade box**. In this decade-inductance box precision inductors are switched in-to or out-of circuit by means of rotary switches. Decade variable inductor is used to carryout experiments and in Inductance (L) meters.

Special types of Inductors

Certain electronic circuits use a special type of Inductor called **Thin-film inductors**. These inductors are thin metal films deposited in the form of a spiral on a ceramic or epoxy base. These are tiny sized and have very low value of inductance.

Copper tube Inductors: At high frequencies, current has a tendency to flow in the skin of the conductor, this is known as **skin effect**. Therefore at high frequency & high power applications hollow copper tube coil is used as inductor instead of solid copper wire.

Variometers: If different radio frequencies are to be received using a single antenna, the electrical length of the antenna will have to be varied, to respond to different wave lengths. Variable inductors used to achieve this are called variometers.

INDUCTANCE MEASURING INSTRUMENTS

Instruments that operate on the principle of wheatstone bridge are used to measure inductance of inductors. These instruments are known as Impedance Bridge, RLC Bridge and so on.

While measuring inductance value using these bridges, an internally generated 1 KHz signal is used for measurement. However an external signal generator may be used to measure the Q of coils at any desired frequency (Q stands for quality-factor: discussed in further lessons).

These instruments can be used to measure inductance values from 1 μH to 1000 H.

Digital Instruments are also available to measure inductance values ranging from 1 μH to 10 H. These Digital meters are simple to operate and are also highly accurate. The meters are commonly known as Digital LC Meters, Digital RLC meters and so on.

Ohmic resistance/DC resistance of INDUCTORS

Inductors are made using generally copper wire. As shown in Fig 11a, any wire will have some amount of ohmic resistance depending upon its length, diameter and resistivity (specific resistance) of the material. This ohmic resistance of an inductor can either be measured using an ohmmeter or calculated using the formula,

$$R = \frac{\rho l}{A}$$

where,

R is the resistance in ohms

ρ is the resistivity of the material in Ohm meters, $\Omega \cdot \text{m}$

l is the length of wire in metres, m

A is the cross-sectional area of the wire in metre², m²

Refer Appendix D for resistivity of substances at room temperature(20°C).

Example: What is the ohmic or DC resistance of a copper wire of 2mm dia and length 100 metres

Cross-section Area =

$$\frac{\pi d^2}{4} = \frac{\pi (2 \times 10^{-3} \text{ m})^2}{4} = 3.14 \times 10^{-6} \text{ m}^2$$

Resistivity, r of copper from Appendix D is

$$1.72 \times 10^{-8} \Omega \cdot \text{m}$$

Therefore,

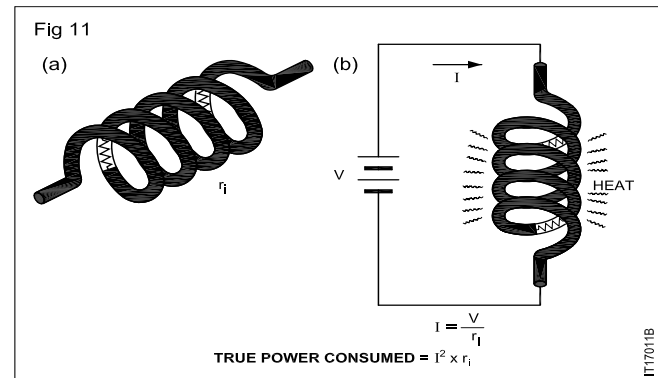
$$R = \frac{\rho l}{A} = \frac{1.72 \times 10^{-8} \Omega \cdot \text{m} \times 100 \text{ m}}{3.14 \times 10^{-6} \text{ m}^2} = 0.55 \Omega$$

The ohmic resistance of the inductor is 0.55 Ω .

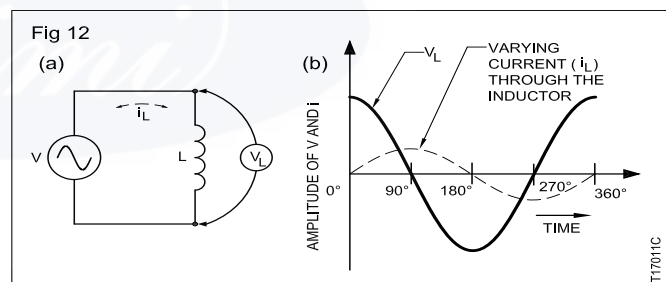
Referring to Fig 11b, because of the DC resistance of the inductor, when current is passed through it, whether DC or AC, some amount of power is consumed. This power is referred as True power given by,

$$\text{True Power} = \text{Current}^2 \times \text{DC resistance} = I^2 R \dots\dots[1]$$

Now consider the effect of only the inductance in a circuit.



When a DC voltage is applied to an inductor, the inductor has its effect only during switching ON and switching OFF of circuit as discussed and illustrated in Fig 2 of this lesson. With steady current in circuit, inductance has no effect. Instead of DC, if a sinusoidal AC current is made to flow through an inductor, as shown in Fig 12a, since the magnitude of sinusoidal current is continuously varying, as shown in Fig 12b, the inductor continuously keep opposing these changes. This continuous opposition is entirely dependent on the induced emf(cemf) in the coil and has nothing to do with opposition due to the DC resistance of the coil. The effective opposition to the flow of alternating current, due to the self induced emf in an inductor is called the Inductive reactance. The unit of inductive reactance is also ohms.



As in a resistive circuit, where opposition to current is given by,

$$R = \frac{V_R}{I_R} \text{ Ohms}$$

where,

V_R = Voltage across the resistor

and I_R = Current through the resistor

similarly, the opposition to current by a pure inductance is given by,

$$X_L = \frac{V_L}{I_L} \text{ Ohms}$$

where,

X_L is the inductive reactance in ohms, Ω

V_L is the voltage across the pure inductor in volts, V

I_L is the current through the inductor in amperes, A

Phase angle relationship between current & voltage

If an inductance of negligible DC resistance is connected across a sinusoidal AC source as shown in Fig 13a, the instantaneous value of current i is given by,

$$i = I_m \sin \omega t, \text{ as shown in Fig 13b} \quad \dots[2]$$

where, I_m is the peak current $\omega = 2\pi f$, where f is the frequency of AC.

The instantaneous voltage across the inductor is given by,

$$v = L \left(\frac{\Delta i}{\Delta t} \right)$$

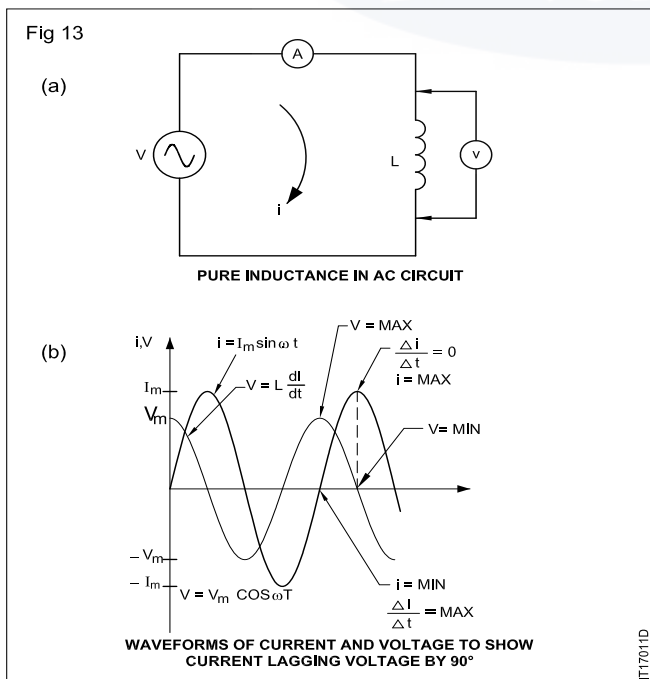
where,

$\frac{\Delta i}{\Delta t}$ is the slope(rate of change) of current waveform.

Referring to Fig 13b, when the value of current i is zero(when it crosses time axis), maximum voltage would be induced across the inductor. This is because the inductor opposes this change $\frac{\Delta i}{\Delta t} = \text{max}$ in current.

So the instantaneous voltage across the inductor(v) will be maximum when the current i is minimum. This is shown in the graph at Fig 13b.

The voltage across the inductor as shown in Fig 13b, is also sinusoidal in shape. In fact, the voltage waveform is called as COSINE curve because, the voltage curve has shifted to the left, or advanced by a quarter of a cycle or 90° . The instantaneous value of voltage across inductor(v) is therefore given by, $v = V_m \cos \omega t$ where, ω is called the angular velocity, given by $2\pi f$.



Since the voltage wave form is advanced by 90° to current wave form, we can say that "in a pure inductive circuit, current lags behind the applied voltage by 90° ".

Factors affecting Inductive reactance

Inductive reactance or opposition to current in an inductor is due to the induced emf in the inductance (v_L). If v_L increases the inductive reactance also increases.

Recall the equation, $v_L = L \left(\frac{\Delta i}{\Delta t} \right)$.

From this equation, it is relevant that, if L is increased, v_L increases and hence inductive reactance also increases.

The current slope $\frac{\Delta i}{\Delta t}$, depends on frequency. If frequency is increased, the time required (Δt) for the current to reach its maximum value decreases and therefore $\frac{\Delta i}{\Delta t}$ increases. How fast $\frac{\Delta i}{\Delta t}$ is changing is given by the angular frequency ω ($\omega = 2\pi f$). Combining these two effects (L and ω), **inductive reactance** X_L is given by,

$$\begin{aligned} X_L &= \omega L \text{ ohms} \\ &= 2\pi f L \text{ ohms} \\ &= 6.28 f L \text{ ohms} \end{aligned} \quad \dots[3]$$

where,

X_L is inductive reactance in ohms, Ω

f is the frequency in hertz, Hz

L is the inductance in henrys, H

ω is the angular velocity in radians per second, rad/sec

Summarizing, **inductive reactance** of given inductor,

- increases as the frequency of the applied voltage increases
- increases if the inductance value of the inductor is increased.

Example: What is the reactance of a 3 H inductor when the frequency of the applied AC is 60 Hz ?

$$X_L = 2\pi f L = 6.28 f L = 6.28 \times 60 \times 3 = \mathbf{1130.4 \Omega}$$

In the above example, if the frequency of AC is increased to 120 Hz, then X_L will be,

$$X_L = 2\pi f L = 6.28 f L = 6.28 \times 120 \times 3 = \mathbf{2260.8 \Omega}$$

In the same example, if the inductance is increased to 6 H with frequency of AC at 60 Hz then X_L will be;

$$X_L = 2\pi f L = 6.28 f L = 6.28 \times 60 \times 6 = \mathbf{2260.8 \Omega}$$

Power consumed by a pure inductor

The power consumed by a pure resistor is given by;

$$P = I^2 R = I.V \quad (V = I.R)$$

The power consumed by a component having both resistive and reactive component is given by $P = V.I.\cos\phi$

where,

V = Voltage applied across the component

I = Current through the component

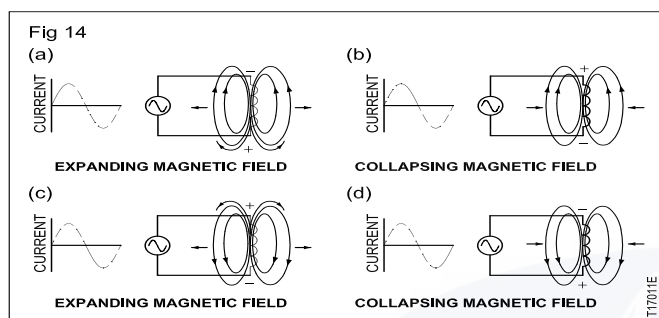
and ϕ = Phase angle difference between V and I

In a pure inductor, as shown in Fig 13b, phase difference between V and I is 90° . Therefore, the power consumed by a pure inductor is $P_{\text{inductor}} = V.I.\cos 90^\circ$

$$\cos 90^\circ = 0 \text{ (Refer trigonometric tables)}$$

$$\text{Therefore } P_{\text{inductor}} = V.I.0 = 0.$$

From energy point of view, we can say that energy is transferred back and forth between the source and the inductor as shown in Fig 14.

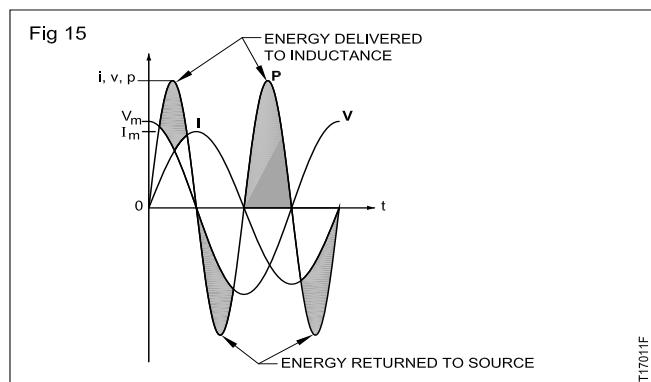


During the quarter cycle, when the current is rising as shown in Fig 14a, energy is taken from the source and converted into magnetic energy & stored in the inductor's field.

During the next quarter cycle, when the current is decreasing as shown in Fig 14b, the field of the inductor is collapsing. The stored energy in the inductor is now converted back to electric energy and returned to the source.

Fig 14(c) & (d) shows the energy stored and energy released during the -ve half cycle of the source.

In Fig 15, the shaded portion of the power curve above and below the zero axis represents the energy delivered to the inductor and energy returned by the inductor during consecutive positive and negative half cycles. This means, if a **Watt meter** is connected in the circuit to measure the power consumed by the inductor, the wattmeter reads zero indicating no power is consumed by the inductor.

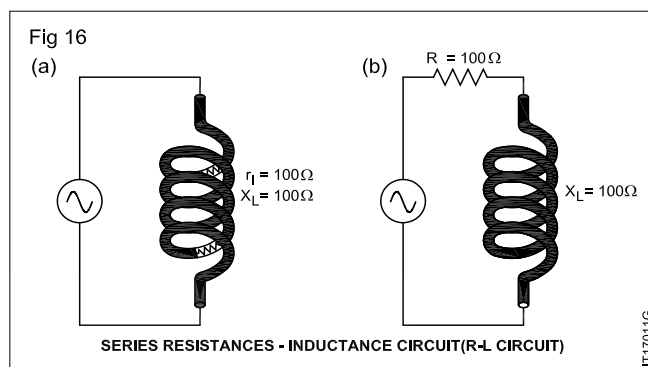


NOTE: Wattmeters are electrical instruments used to measure DC or AC power directly.

Practical inductors

So far ideal inductors with zero DC resistance, (generally called ohmic resistance) were considered. In practice there is always some ohmic resistance associated with inductors. This is because, inductors are made of copper wire, which has finite ohmic resistance.

Consider a practical inductor with some finite ohmic resistance as shown in Fig 16a. This inductor has ohmic resistance (r_l) of 100Ω and an inductive reactance of 100Ω . Fig 16a can be rewritten as in Fig 16b, as if there is a 100Ω resistor in series with an ideal inductor having inductive reactance 100Ω .



Series R - L circuit

Referring to Fig 17a, which is same as Fig 16b, the circuit current (I) is limited by both the ohmic resistance R and inductive reactance X_L . Each has its own series voltage drops IR and IX_L . Here the circuit current is labeled as I , instead of I_L , because current I flows through both the series components R and L .

Graph at Fig 17b shows the instantaneous values of i , V_R , V_L and the source voltage V_S . A vector diagram of V_R , V_L and I is shown in Fig 17c. The vector diagram at Fig 17c, shows that the current I lags behind the voltage V_L by 90° . But the current I is in phase with V_R .

From the graph at Fig 17b, V_R is maximum ($100V$) when V_L is minimum and vice versa. This is again because of the phase difference. Because of this the series voltage drops V_R and V_L cannot be added arithmetically to get the applied source voltage V_S . The method of adding V_R and V_L is shown in Fig 18.

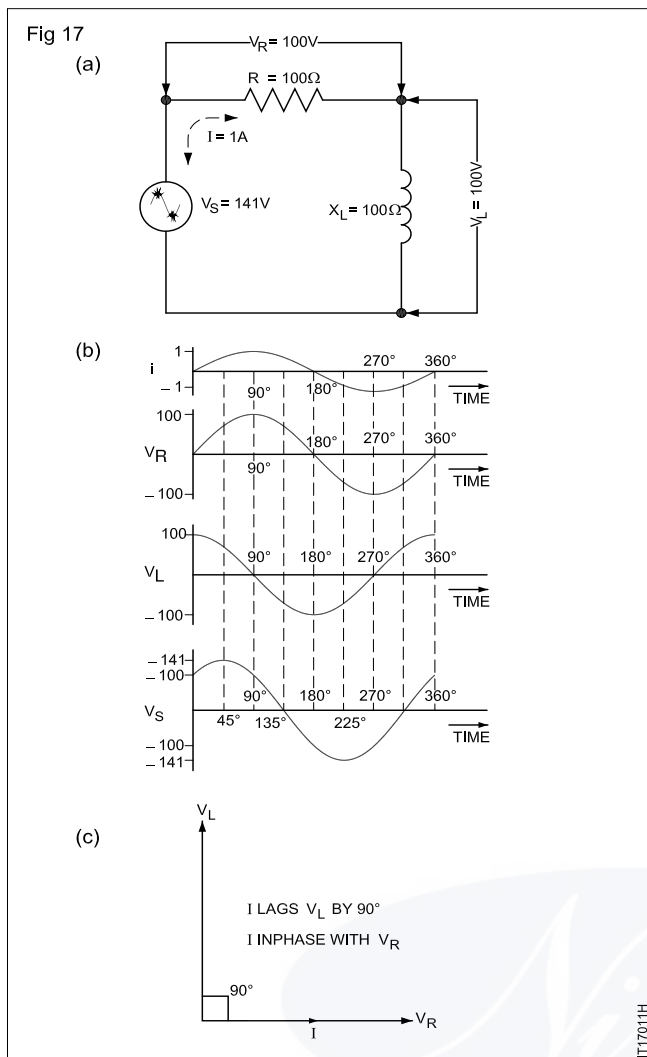
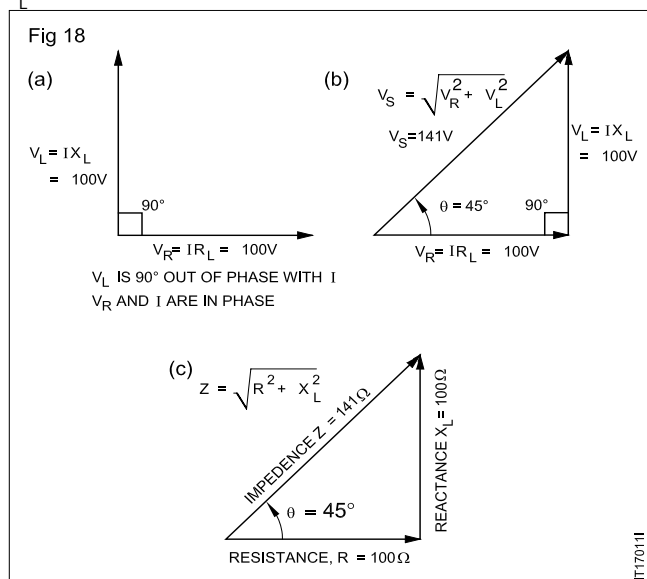


Fig 18a shows the vectorial addition method to get V_S knowing V_R and V_L . Fig 18b is a redrawn form of Fig 18a to find V_S from V_R and V_L .

Fig 18c gives the total resultant opposition to current flow due to R and X_L . This total resultant opposition due to resistance R and inductive reactance X_L is called **Impedance** in ohms with the symbol Z . The impedance Z , takes into account the phase relationship between R and X_L .



The impedance Z of the circuit given at Fig 18c is,

$$Z = \sqrt{R^2 + X_L^2} \quad \dots[4]$$

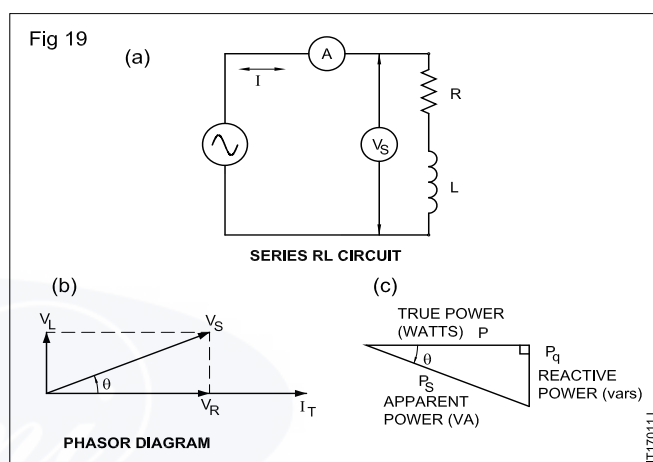
$$= \sqrt{100^2 + 100^2} = 141\Omega$$

Power consumed in a Resistance - Inductance Circuit (R-L circuit)

The total power consumed in a R-L circuit will contain the **True power** due to the pure resistive component and the **Reactive power** due to the reactive component.

The power consumed by a reactive component is referred to as reactive power.

In the series R-L circuit at Fig 19a, *Apparent power* P_s is the vectorial sum of True power (P) and the reactive power (P_q) as shown in Fig 19c.



Apparent power (P_s) is also given by the product of source voltage V_S and total circuit current I_t .

Apparent power = $V_s \times I_t$ in volt-amperes, VA ...[5]

To distinguish from reactive power and apparent power, the power dissipated in a resistor in the form of heat (or in any other form), the term *Real power* or *True power* is used.

True power = $V_R \times I_R$ watts, W

In a series circuit, since the current is same through all components, I_R is same as I_t . Therefore in Fig 19 the True power consumed by the resistor is given by

True power (P) = $V_R \times I_T$ W

Assuming that the internal resistance (r_i) of an inductor is negligible when compared with the high value of a series resistor (r_s), the power consumed by an inductor can be taken as zero. As explained earlier, the power delivered to inductor in one quarter, is returned in the other quarter. However the source should capable of delivering power for the first quarter of the cycle, even though this power is returned during the next quarter cycle. This stored or transferred power is called reactive power, P_q . The subscript q represents quadrature power.

The reactive power of a pure inductor is given by

Reactive power (P_q) = $V_L \times I_L$ volt-amperes reactive or vars

From Fig 19b, in a series R – L circuit source voltage, V_s is equal to,

$$V_s = \sqrt{V_R^2 + V_L^2} \quad \dots[6]$$

$$\text{Therefore, } V_s I_t = \sqrt{(V_R I_t)^2 + (V_L I_t)^2}$$

$$\text{This is nothing but, } P_s = \sqrt{P^2 + P_q^2} \text{ volt-amperes (VA)} \quad \dots[7]$$

This relationship is represented in what is known as **Power triangle** as shown in Fig 19c.

From the power triangle it can be seen that, due to the presence of both inductance and resistance, the circuit current will lag the source voltage by an angle ϕ which is less than 90° . This angle ϕ is commonly called as the phase angle or power factor $\cos\phi$ of the circuit.

From Fig 19c, the power factor of the circuit is,

$$\cos \phi = \frac{\text{True power}}{\text{Apparent power}}.$$

In a right angle triangle, $\cos \phi$ is given by the ratio of adjacent side to hypotenuse.

From Fig 19b, the power factor $\cos\phi$, can be found out as follows:

$$\cos \phi = \frac{V_R}{V_s} = \frac{I_R}{I_Z}$$

Since the current I through all the components in a series circuit is same, the above equation can be rewritten as

$$\cos \phi = \frac{R}{Z} \quad \dots[8]$$

Quality factor - Q of coil

At high frequencies, how useful is a coil is not only judged by its inductance, but also by the ratio of its inductive reactance to its internal DC resistance of the coil. This ratio is called the Quality factor or merit or Q of the coil.

$$Q \text{ of a coil is given by, } Q = \frac{X_L}{R_i}$$

where,

X_L is the reactance of the coil in ohms

R_i is the internal Resistance of the coil in ohms

Since X_L and R_i have the same units of measure, Q has no unit.

The Q of a coil can be defined as the ability of a coil to produce self-induced voltage. The Q factor of a coil can also be defined as the capability of the coil to store energy. Hence the Q factor of a coil is also known as the *storage factor*.

If Q of a coil is 200, it means, that the X_L of the coil is 200 times more than its R_i . Q of coils range from less than 10 for a low Q coil up-to 1000 for a high Q coils. R.F coils have Q in the range from 30 to 300.

CHART – 1
TYPES OF INDUCTORS

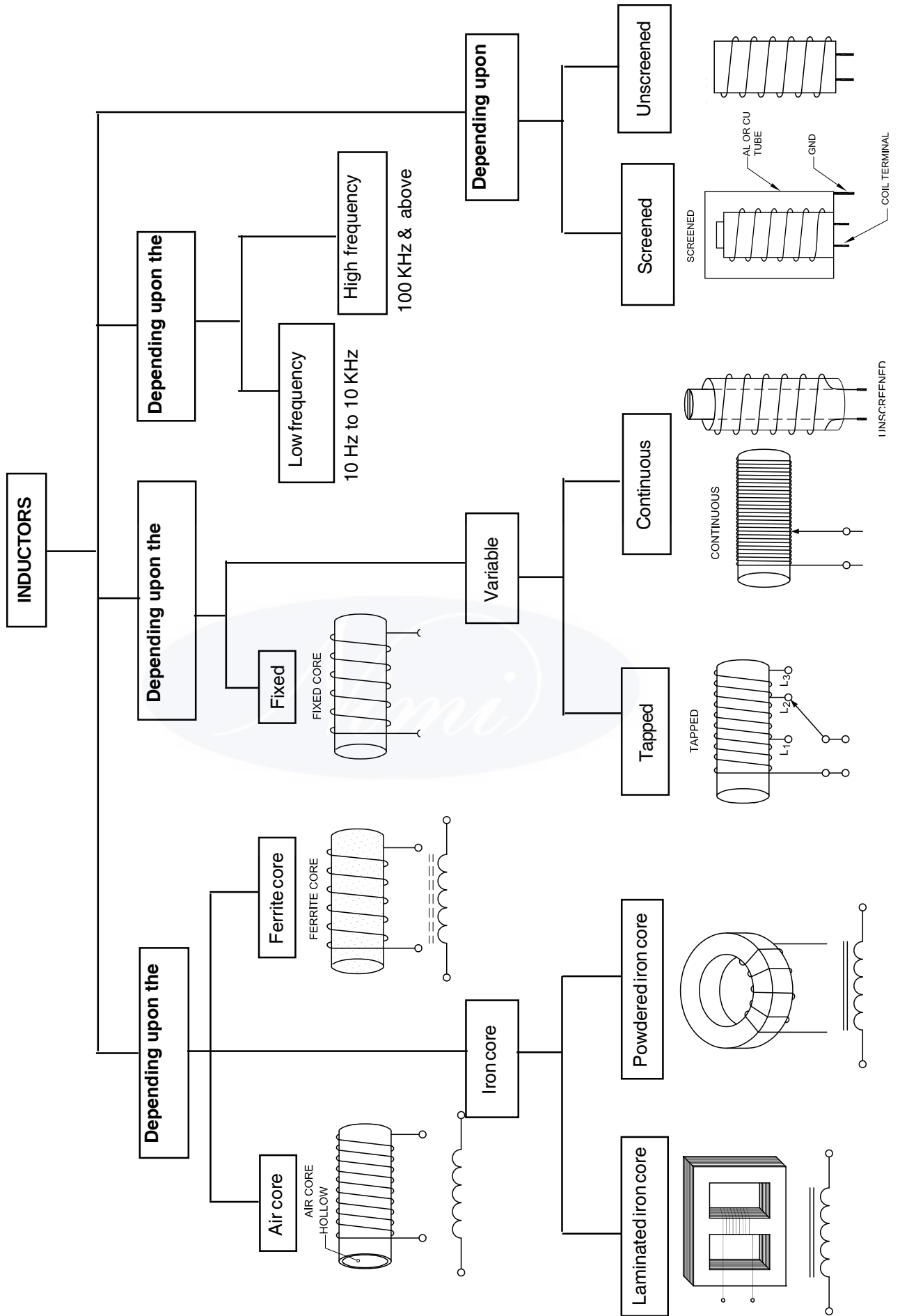
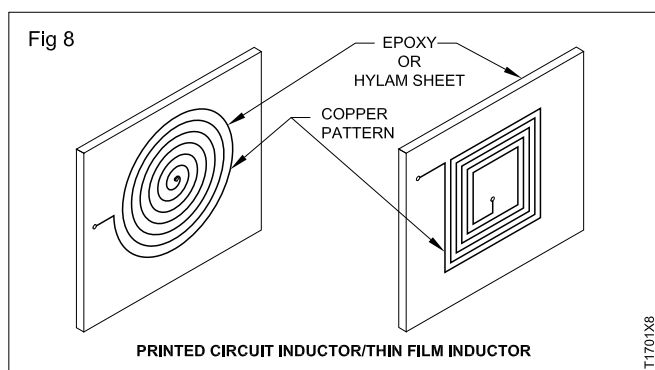
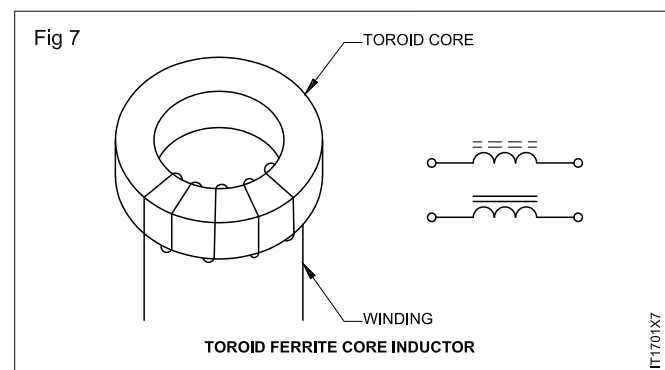
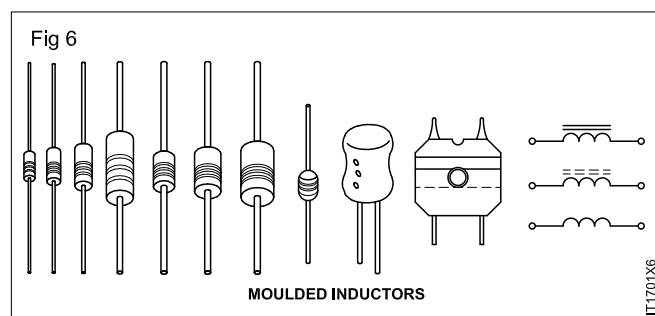
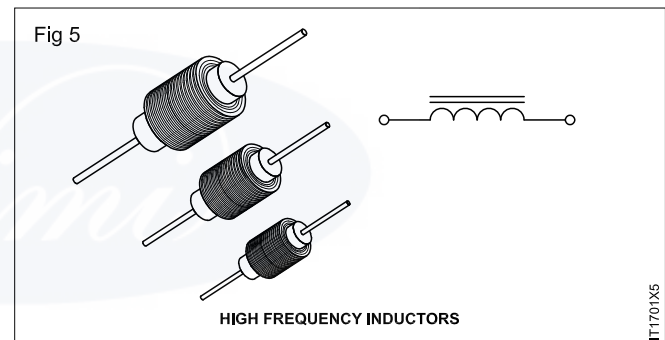
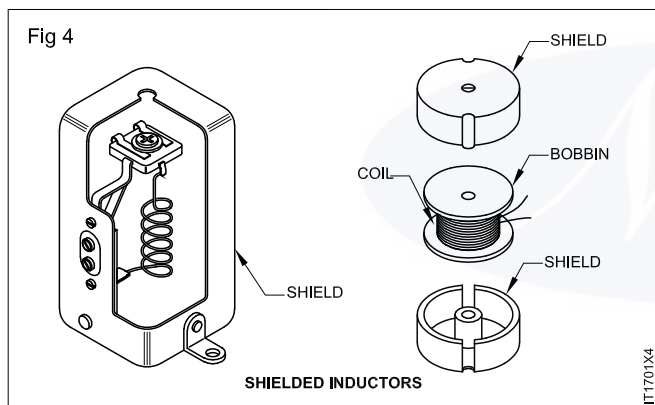
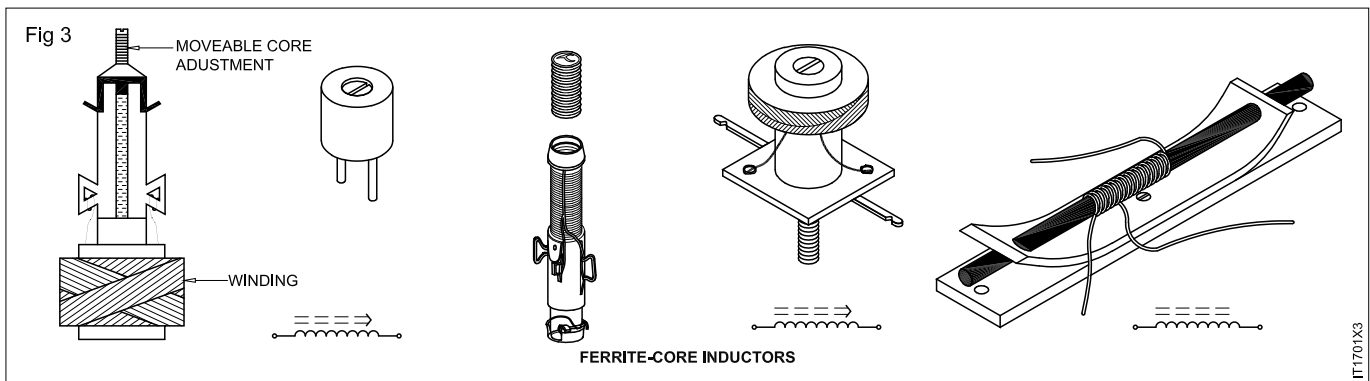
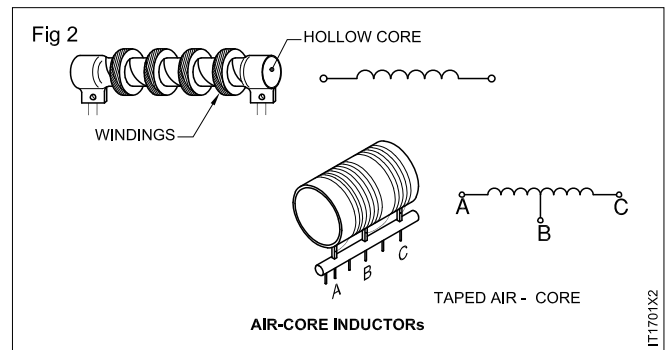
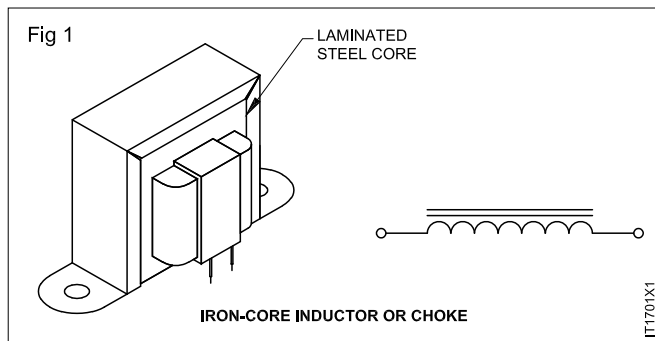


CHART - 2 PHYSICAL APPEARANCE OF DIFFERENT TYPES OF INDUCTORS

IRON CORE INDUCTOR OR CHOKE



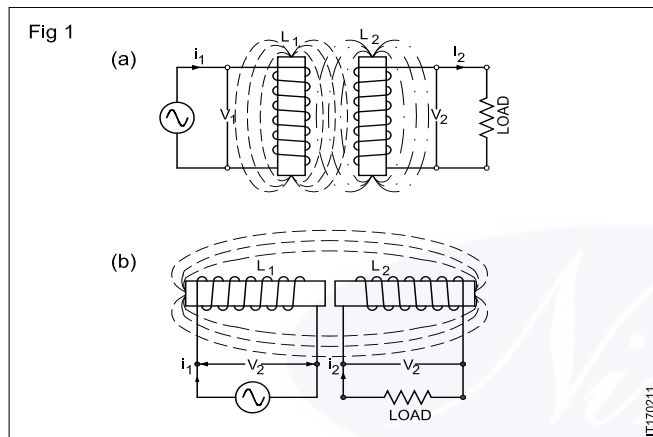
Concept of Self & Mutual Inductance

Objectives: At the end of this lesson you shall be able to

- explain the property of mutual induction
- explain the term coefficient of coupling
- find the value of inductors in series without mutual coupling
- find the value of inductors in series with mutual coupling
- find the value of inductors in parallel
- state method of finding mutual inductance experimentally.

Mutual Inductance (M)

When two inductors L_1 and L_2 are placed side by side close to each other shown in Fig 1a or Fig 1b, although the two coils are not electrically connected, the two coils are said to be magnetically inter-coupled.



The changing current i_1 in coil L_1 not only self induces an emf (V_1) in L_1 , but also causes a voltage (V_2) to be induced in L_2 . The voltage V_2 induced in L_2 causes a current i_2 that sets-up its own changing flux around L_2 . This in turn, not only self induces a voltage in L_2 , but also induces an additional voltage in L_1 . That is, a changing current in one coil will induce an emf in other nearby coil. This effect is known as **mutual induction**.

The two coils L_1 and L_2 of Fig 1, are said to have a mutual inductance (M), in addition to their own self-inductances (L).

Mutual inductance, like self-inductance, is also measured in units of Henrys. The definition is given below;

Two coils are said to have a mutual inductance of 1 Henry, when a current changing at the rate of 1 Amp/sec in one coil induces an emf of 1V in the other coil.

Coefficient of coupling

The amount of mutual inductance (M) between two coils depend upon, the self inductance of each coil and the amount of mutual flux between the two coils.

The amount of mutual flux, that links both coils is dependent on the physical placement of the two coils. This is indicated

by the term **Coefficient of coupling, k** .

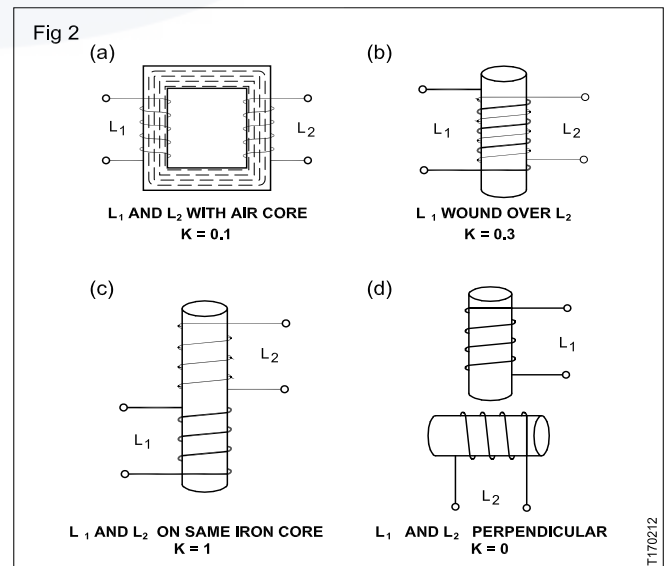
Coefficient of coupling k , between the two coils is given by,

$$k = \frac{\text{Mutual flux between two coils } \phi_m, \text{ in Webers}}{\text{Total flux set up by one coil, in Webers}}$$

Maximum value of k can be 1. This occurs when all the flux (ϕ) set-up by one coil is linking with the other coil. For example; when both the coils are wound as shown in Fig 2a, almost all the flux set-up in one coil is interacting with the other coil. In other words there is very little or zero leakage of flux. In such cases k is practically equal to 1. This condition of $k=1$ is also known as **tight coupling**.

In Fig 2b, if only 30% of the flux set-up by coil 1, links with coil 2, the coefficient of coupling is only 0.3.

In Fig 2c and Fig 2d where the coils are placed far apart or when the two coils are placed perpendicular to one another, the coupling is minimum and will be close to zero.



It can be shown that mutual inductance (M) between the given two coils L_1 and L_2 can be found out using the formula,

$$M = k \sqrt{L_1 \cdot L_2} \text{ Henrys.}$$

Where,

k is the coefficient of coupling which has no units

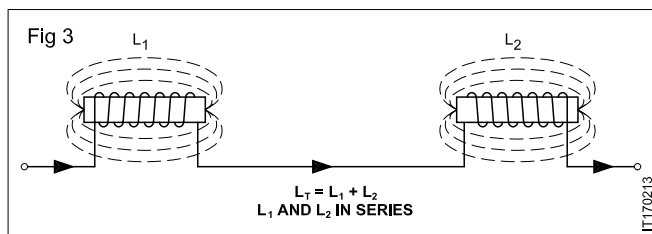
L_1 and L_2 are inductance values, in henrys

M is the mutual inductance, in henrys

INDUCTORS IN SERIES

In order to obtain a desired value of inductance, inductors can be connected either in series or in parallel.

Fig 3 shows two inductances connected in series. The spacing between the inductors are large enough so that there exists no mutual inductance between the two coils. Hence in Fig 3 $k=0$. In Fig 3, since the direction of current is same through both coils, the self-induced voltages are additive. Therefore the total inductance of such series connection is given by,



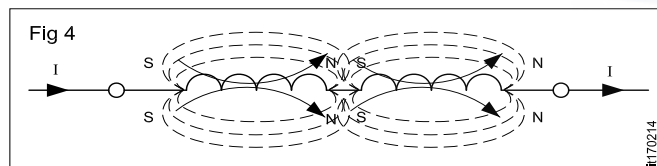
$$L_T = L_1 + L_2 + L_3 + \dots + L_n \text{ Henrys (H)}$$

where, L_T is the total inductance across end terminals.

$L_1, L_2, \dots, L_n$ are individual inductance values.

Series coils with mutual inductance

Unlike in Fig 3, when two inductors L_1 and L_2 are connected in series close to each other as in Fig 4 the total inductance (L_T) will be larger than just the sum of L_1 and L_2 . How much larger will this be depends on the mutual inductance M.



In general, the total inductance of two series-connected coils, with mutual inductance M is given by;

$$L_T = L_1 + L_2 \pm 2M$$

Whether it will be $+2M$ or $-2M$ depends on, whether the inductors are connected in series-aiding as shown in Fig 5a or in series-opposing as shown in Fig 5b.

Dot notation

Whether two coils are connected series-aiding or series-opposing, it is often indicated by using *dot notation* as shown in Fig 6. When current enters both dots or leave both dots as shown in Fig 6a the mutual inductance is additive.

When the current enters one dot and leaves the other dot, as shown in Fig 6b, the mutual inductance is subtractive. In other words the dots indicate the in-phase ends of each other.

Fig 5

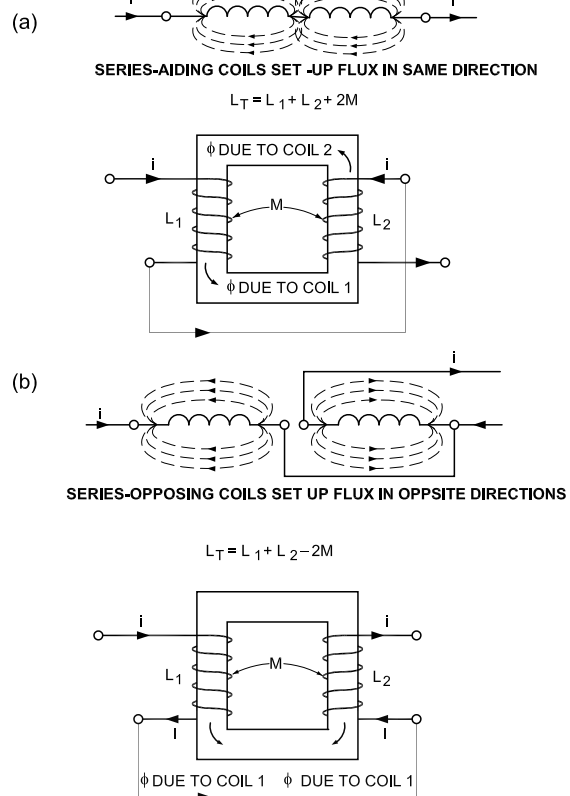
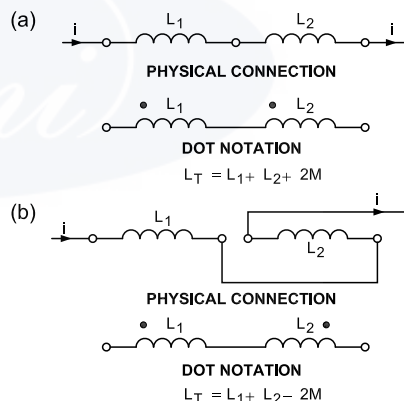


Fig 6



Dot notation is used in transformers and loud speakers to show correct phasing.

Details of dot notation in transformers and speakers will be discussed in further lessons/units.

Example (1): Two inductors of 10H and 15H are placed at right angles to each other. Calculate the total inductance, when the coils are connected in series.

SOLUTION

Since the inductors are perpendicular to each other, mutual inductance (M) is zero.

$$\text{Therefore } L_T = L_1 + L_2 = 10\text{H} + 15\text{H} = 25\text{H}$$

knowing values of L_1 , L_2 and M , k can be found out using the formula.

$$M = k \sqrt{L_1 \cdot L_2} \quad \text{or} \quad k = \frac{M}{\sqrt{L_1 \cdot L_2}}$$

Example (2): If the above two inductors are placed parallel to each other as shown on a common core in Fig 5 with a mutual inductance of 4H and connected in series, calculate,

- L_T of coils when connected series-aiding (Fig 5a).
- L_T of coils when connected series-opposing (Fig 5b).
- Coefficient of coupling k .

SOLUTION

(a) $L_T = L_1 + L_2 + 2M = 10H + 15H + 2 \times 4H = 33H$

(b) $L_T = L_1 + L_2 - 2M = 10 + 15 - 2 \times 4H = 17H$

(c) $M = k \sqrt{L_1 \cdot L_2}$

or, $k = \frac{M}{\sqrt{L_1 L_2}} = \frac{4}{\sqrt{10 \times 15}} = 0.33$

INDUCTORS IN PARALLEL

When two inductors are connected in parallel as shown in Fig 7a, the total inductance is given by,

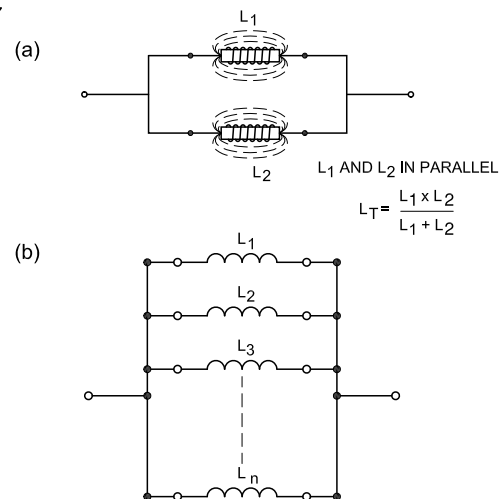
$$\frac{1}{L_T} = \frac{1}{L_1} + \frac{1}{L_2} \quad \text{Henrys (H)}$$

$$L_T = \frac{L_1 \cdot L_2}{L_1 + L_2} \quad \text{Henrys}$$

This is similar to finding effective value of resistors in parallel

In general, when several inductors are connected in parallel as shown in Fig 7b the total inductance is given by

Fig 7



$$\frac{1}{L_T} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + \dots + \frac{1}{L_n} \quad \text{Henrys (H)}$$

Measurement of mutual inductance

The mutual inductance (M) and coefficient of coupling (k) between two coils can be measured as follows;

First : Measure the total inductance with fields aiding,

$$\text{i.e., } L_{Taid} = L_1 + L_2 + 2M \quad \dots\dots[1]$$

Second: Measure the total inductance with fields opposing,

$$\text{i.e., } L_{Topp} = L_1 + L_2 - 2M \quad \dots\dots[2]$$

Solving equations 1 and 2 we get,

$$L_{Taid} - L_{Topp} = 4M \quad \text{or}$$

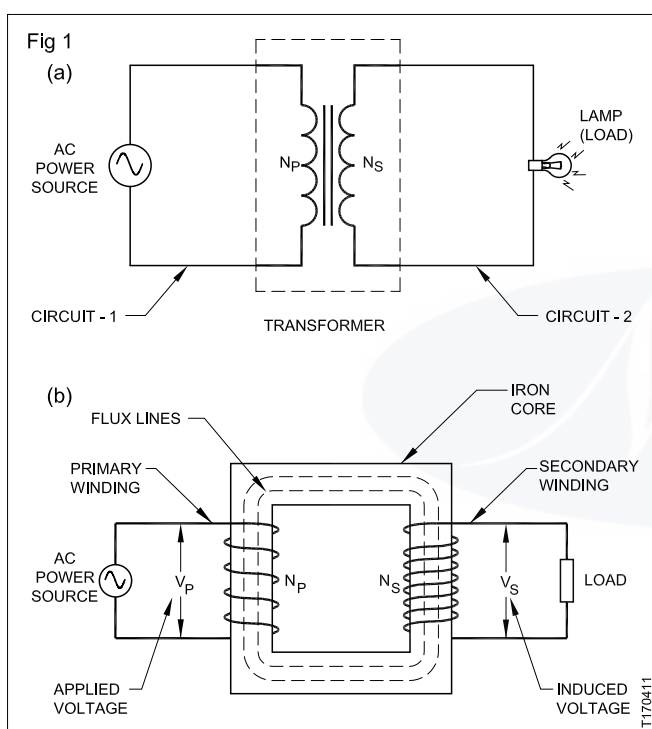
$$M = \frac{L_{Taid} - L_{Topp}}{4}$$

Transformers and its Types

Objectives: At the end of this lesson you shall be able to

- state the basic function of a transformer
- state and explain important terms used with transformers
- list the different classifications of transformers
- list the common application of different types of transformers
- state the difference between an ideal and practical transformer
- define the term efficiency of a transformer and list out the losses
- explain how the losses in a transformer can be minimised
- explain how OC and SC tests reveal losses in transformers.

Transformer is an electrical device used to transfer electric energy from one AC circuit to another circuit by magnetic coupling as shown in Fig 1a.



A transformer essentially consists of two coils of insulated conducting material, generally copper. These coils are wound on a core made of iron or ferrite as shown in Fig 1b. These coils are so arranged that magnetic flux developed in one coil will link with the other coil. Hence, mutual inductance exists between the two coils with tight-coupling ($k=1$). A change in current through one coil (say N_p) induces a voltage in the other coil (say N_s). The magnitude of induced voltage in the secondary winding depends on the number of turns of the coils and on how tight the magnetic coupling (k) is, between the two coils.

In a transformer, as shown in Fig 1b, the coil or the winding to which electrical energy is given from an ac power source is called the **primary winding**. In Fig 1 this coil is marked N_p .

The second coil to which, energy from the primary winding is coupled magnetically is called the secondary winding (N_s in Fig 1b). If a load, say a lamp or a resistor, is connected across the secondary winding, current flows through the load although there is no direct AC power source connected to it.

Hence, transformers can be defined as devices that make use of the principle of mutual induction, in transferring electrical energy from one ac circuit to another circuit without direct electrical connection.

It is important to note that transformers cannot transfer DC energy from primary winding to secondary winding, because, a DC current cannot produce changing magnetic field and hence cannot develop induced voltage.

Important terms used with iron-core transformers are explained below;

1. Turns Ratio of a transformer

The ratio of the number of turns of coil in the primary (N_p) to the number of turns of coil in the secondary (N_s) is called the turns ratio of the transformer.

$$\text{Turns ratio} = \frac{N_p}{N_s}$$

For example, 1000 turns in the primary and 100 turns in the secondary gives a turns ratio of 1000/100, or 10:1 which is stated as ten-to-one turns ratio.

2. Voltage Ratio of a transformer

The ratio of voltage across the primary winding (V_p) to the voltage available across the secondary winding (V_s) is called the voltage ratio of the transformer.

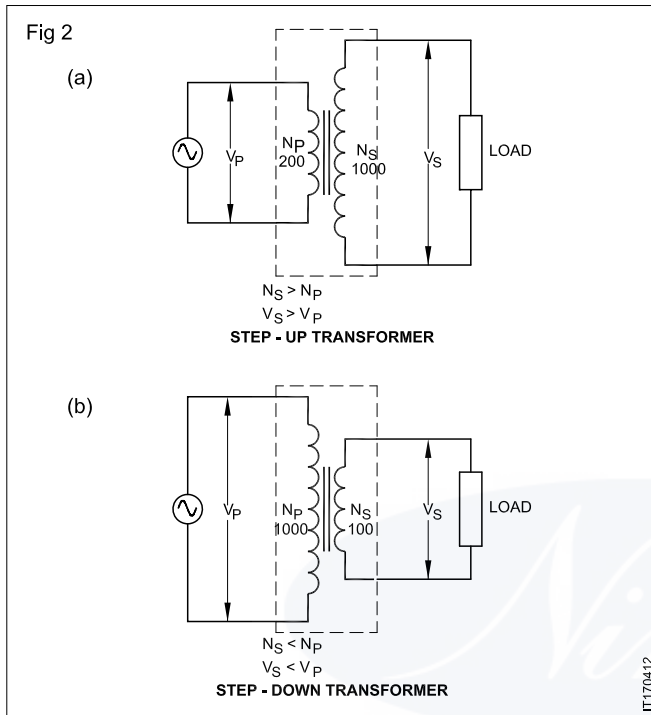
$$\text{Voltage ratio} = \frac{V_p}{V_s}$$

When coefficient of mutual coupling (k) between primary and secondary winding is 1, the voltage induced per turn of the secondary winding is the same as the self-induced voltage per turn in the primary winding. The total voltage appearing across the secondary winding depends on the number of turns of secondary winding.

Therefore, the voltage ratio is in the same proportion as the turn ratio:

$$\frac{V_p}{V_s} = \frac{N_p}{N_s}$$

This means, if the secondary winding has more turns than the primary winding ($N_s > N_p$) then, the secondary voltage will be higher than the primary voltage. In other words, in such a condition the primary voltage is said to be raised or stepped-up. Such transformers are called **STEP-UP TRANSFORMERS** as shown in Fig 2a.



Example: As shown in Fig 2a, a transformer has 200 turns of N_p and 1000 turns of N_s , its turns ratio will be,

$$\text{Turns ratio} = \frac{N_p}{N_s} = \frac{200}{1000} = 1 : 5$$

For this transformer, If the applied AC primary voltage (V_p) is $110 V_{rms}$, the secondary voltage will be stepped up in the same ratio as that of turns ratio. Hence, the secondary voltage will be twice the primary voltage, i.e, $5 \times 110 = 550 V_{rms}$.

On the other hand, when the secondary winding has less number of turns than the primary winding, the primary voltage is said to be lowered or stepped-down. Such transformers are called **STEP-DOWN TRANSFORMERS** as shown in Fig 2b.

Example: As shown in Fig 2b a power transformer has 1000 turns of N_p and 100 turns of N_s . What is the turns ratio? How much is the secondary voltage V_s when a primary voltage is 240 V ?

SOLUTION:

The turns ratio is $1000/100$, or 10:1. Hence, secondary voltage will be stepped down by a factor of 1/10, making V_s equal to $240/10$ or 24 Volts.

3. Secondary Current in transformers

The amount of current in secondary winding of a transformer will be equal to, by ohms's law, the secondary voltage V_s divided by the load resistance in the secondary circuit.

Referring to Fig 3, If the value of load (R_L) is 10 ohms and if the secondary winding coil resistance is negligible, then, secondary current I_s will be,

$$I_s = \frac{V_s}{R_L} = \frac{12 V}{10 \Omega} = 1.2 \text{ Amperes}$$

4. Power in the Secondary of transformer

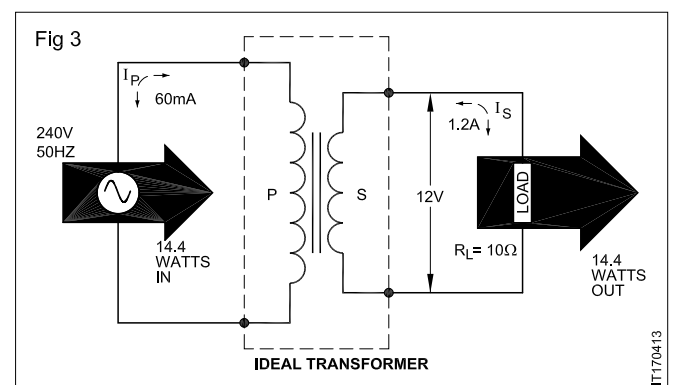
Referring to Fig 3, the power dissipated by load R_L in the secondary (P_s) is given by;

$$P_s = (I_s^2 \times R_L) \text{ or } (V_s \times I_s) \\ = ((1.2)^2 \times 10) \text{ or } (12 \times 1.2) = 14.4 \text{ watts}$$

Referring to the primary of Fig 3, the power consumed at the primary (P_p) is given by,

$$P_p = V_p \times I_p \\ = 240V \times 60mA \\ = \frac{240 \times 60}{1000} = 14.4 \text{ watts.}$$

From the above calculations it can be seen that the power consumed in primary is same as the power consumed in secondary ($P_s = P_p$). Note that a current of 60mA at 240 volts in primary results in 1.2Amps at 12V in secondary. Thus in an ideal transformer (having no energy loss) the power consumed in the secondary is equal to the power consumed in the primary of the transformer. This actually means that, the power consumed by the secondary is fully supplied by the primary as the secondary has no direct electrical connection with the supply source (240V, 50Hz in Fig 3).



Current Ratio of transformer

Referring to Fig 3, in an ideal transformer the power consumed in the secondary equals the power in the primary. Mathematically it is given by,

$$P_s = P_p \text{ or } V_s \times I_s = V_p \times I_p$$

This equation can be rearranged as,

$$\frac{I_s}{I_p} = \frac{V_p}{V_s}$$

Since $\frac{V_p}{V_s} = \frac{N_p}{N_s}$, $\frac{I_s}{I_p} = \frac{N_p}{N_s}$

Summarizing $\frac{N_p}{N_s} = \frac{V_p}{V_s} = \frac{I_s}{I_p}$ [1]

As it can be seen, the current ratio I_s/I_p , is the inverse of the voltage ratio V_p/V_s . This means voltage step-up in the secondary means current step-down, and vice versa.

Example: A voltage step-up transformer with a 1:5 turns ratio has 220V across 2200W in the secondary. (a) How much is I_s ? (b) Calculate the value of I_p .

SOLUTION:

(a) $I_s = \frac{V_s}{R_L} = \frac{200 \text{ V}}{2200 \Omega} = 0.1 \text{ A}$

(b) With a turns ratio of 1:5,

$$\frac{N_p}{N_s} = \frac{1}{5}$$

$$\frac{I_s}{I_p} = \frac{N_p}{N_s} = \frac{1}{5}$$

$$I_p = 5 \times I_s = 5 \times 0.1 = 0.5 \text{ A}$$

Example: A transformer with a 20:1 voltage step-down ratio has 12V across 1 W in the secondary. (a) How much is I_s ? (b) How much is I_p ?

SOLUTION:

(a) $I_s = \frac{V_s}{R_L} = \frac{12 \text{ V}}{1 \Omega} = 12 \text{ A}$

(b)

$$\frac{N_p}{N_s} = \frac{20}{1}$$

$$\frac{I_s}{I_p} = \frac{N_p}{N_s} = \frac{20}{1}$$

$$I_p = \frac{I_s}{20} = \frac{12 \text{ A}}{20} = 0.6 \text{ A}$$

As a check for your calculations, remember that the side with the higher voltage always has the lower current.

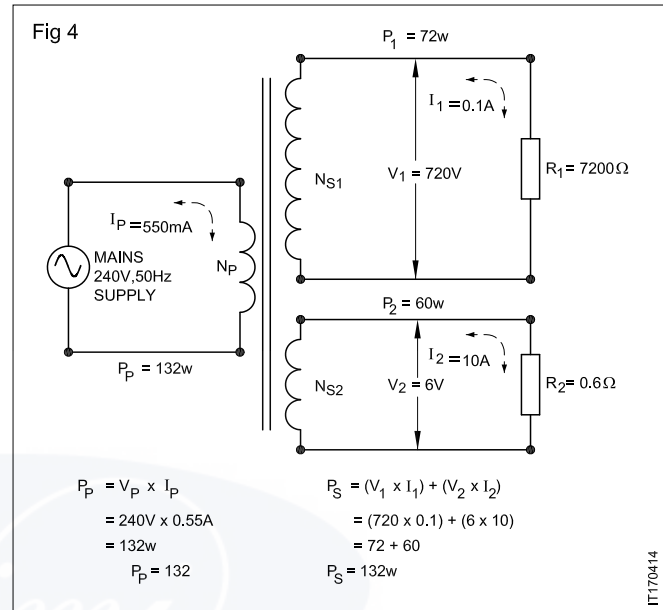
Power in Transformers with more than one secondary winding

A transformer need not have only one secondary winding as shown in Figures 1 to 3. Transformers can have one, two, or more number of secondary windings. Each

secondary winding will have induced voltage in proportion to its turns ratio with the common primary winding. Fig 4 shows a transformer with two secondary windings.

Referring to Fig 4, the secondary winding N_{s1} provides voltage step-up and has a step-up of 3:1 providing 720 V. The other winding N_{s2} provides voltage step-down, with a ratio of 40:1 providing 6 V.

If the first secondary circuit consumes 72 watts and the second secondary circuit consumes 60 watts, then the power supplied by the 240 V source in the primary is $72 + 60 = 132 \text{ W}$.

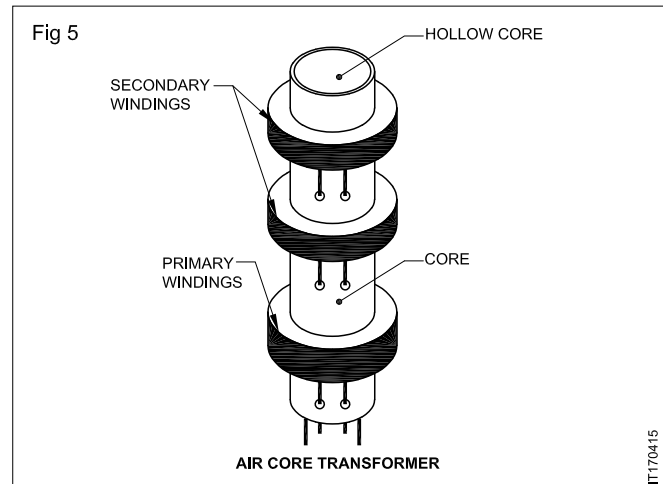


CLASSIFICATION OF TRANSFORMERS

1 Classification based on the type of Core Material used

Transformers can be classified according to the type of material used for the core as;

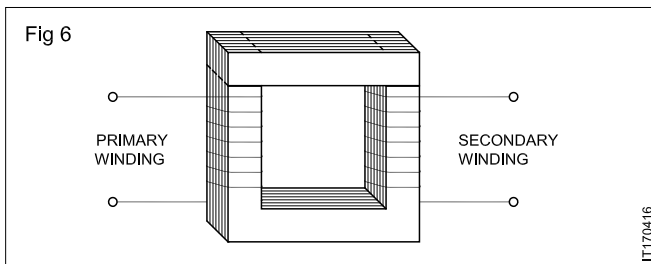
(a) **Air core transformers** : As shown in Fig 5, air core transformers consists of a hollow non magnetic core, made of paper or plastic over which the primary and secondary windings are wound. These transformers will have values of k less than 1. Air core transformers are



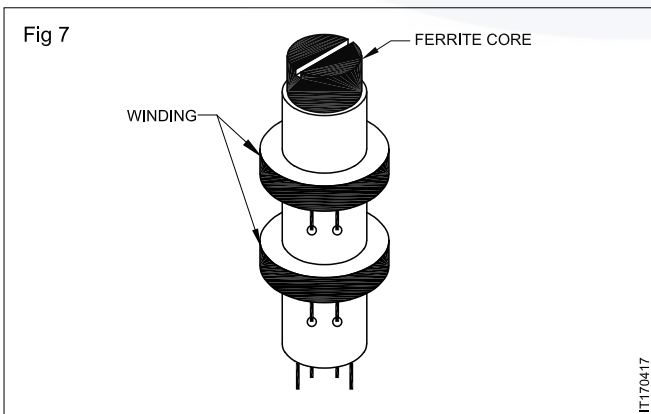
generally used in high frequency applications because these will have no *iron-loss* as there is no magnetic core material.

Iron-loss is a type of transformer loss due to core material. Transformer losses are discussed in detail in further lessons.

- (b) **Iron core transformers:** Fig 6 shows a laminated iron-core transformer. These transformers have stacked laminated sheets of silicon steel over which the windings are wound. This is the most common type of transformer used with mains power supply (240V, 50Hz). In these transformers, since the core is a magnetic material and due to the shape of the core, the value of k is almost equal to 1.



- (c) **Ferrite core transformers:** As shown in Fig 7, these transformers have Ferrite material as its core. In most cases, the primary and secondary windings are wound on a hollow plastic core and the ferrite material is then inserted into the hollow core. These transformers are used in high frequency to very high frequency applications as they have the advantage of introducing minimum losses. In these transformer, the position of core can be changed, thus changing the value of M .



2 Classification based on the shape of core and relative position of primary and secondary windings:

- (a) **Core type transformers:** Refer Fig 1 of Chart 1 given at the end of this lesson. In Core type of transformer, the primary and secondary windings are on two separate sections/limb of core. Core type transformers are less frequently used as their efficiency is low because the magnetic flux spreads out reducing the number of useful flux lines. The shape of core stampings used for these transformers is shown in Fig 1c.

- (b) **Shell type transformers:** Refer Fig 2 of Chart 1. In this type, both the primary and the secondary windings are wound on the same section/limb of the core. As the portion of the core surrounds the two windings, almost all the flux is confined to the core of the transformer. Shell type transformers have a higher efficiency as compared to core type transformers. These are widely used as voltage and power transformers. The shape of core stamping used for these transformers is shown in Fig 2c.

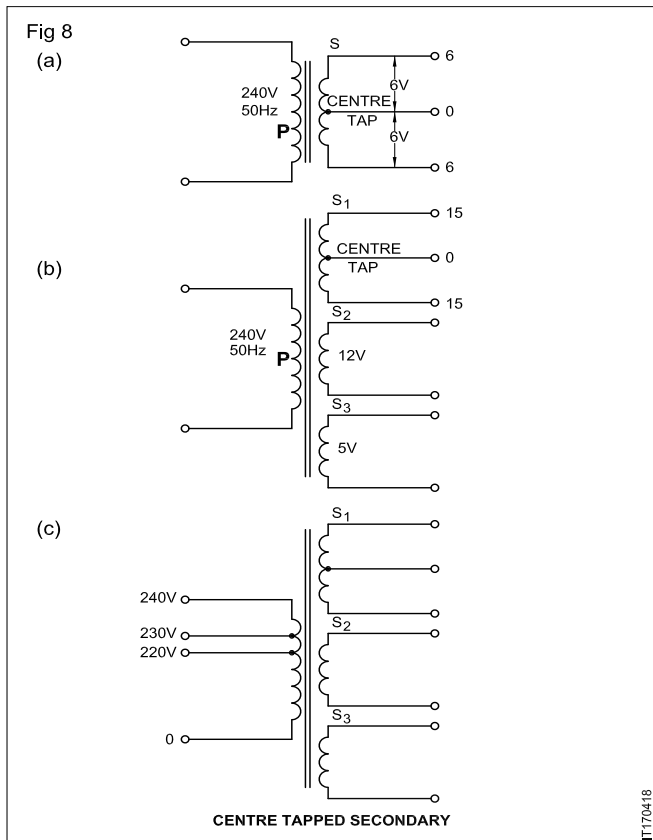
- (c) **Ring type transformers:** Refer Fig 3 of Chart 1. In this, the core is made up of circular or semicircular laminations as shown in Fig 3c. These are stacked and clamped together to form a ring. The primary and secondary windings are then wound on the ring. The disadvantage of this type of construction is the difficulty involved in winding the primary and secondary coils. Ring type transformers are generally used as instrument transformers for measurement of high voltage and current.

3 Classification based on the Transformation ratio:

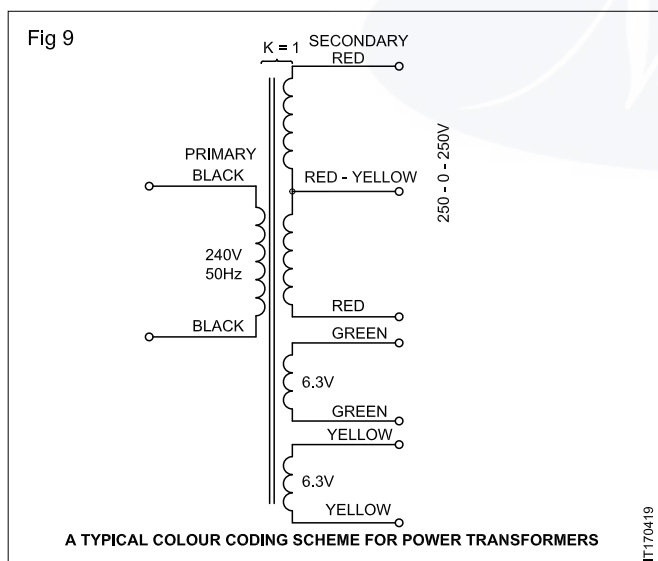
- a) **Step-up Transformers:** Transformers in which, the induced secondary voltage is higher than the source voltage given at primary are called *step-up transformers*.
- b) **Step-down Transformers:** Transformers in which, the induced secondary voltage is lower than the source voltage given at primary are called *step-down transformers*.
- c) **Isolation transformers:** Transformers in which, the induced secondary voltage is same as that of the source voltage given at primary are called *one-to-one* or *isolation transformers*. In these transformers the number of turns in the secondary will be equal to the number of turns in the primary making the turns ratio equal to 1.

4 Classification based on the operating frequency:

- a) **Mains transformer:** Refer Fig 4 of Chart 1. These are basically, iron-core shell type transformers. These transformers form the link between AC mains source and other devices requiring AC or DC power. As an example, radio receivers need low voltage DC supply. To use a radio receiver from AC mains, the mains AC is stepped-down, AC is converted to DC (rectification) and fed to the radio receiver. The secondary winding of these transformers may have a *centre tap* as shown in Fig 8a or may have more than one secondary windings as shown in Fig 8b. These transformers may also have more than two terminals at primary winding as shown in Fig 8c to accommodate for different AC mains levels. Tapped primary also allows changes in the secondary-primary turns ratio. All the power transformers are generally designed to work at mains supply frequency (50 Hz).



Power transformers use colour coding scheme to identify the primary and secondary windings. One such scheme is shown in Fig 9.



b) **Audio frequency (AF) transformers:** Refer Fig 5 of Chart 1. These AF transformers look similar to a mains voltage transformer but they are very small in size comparatively. Most AF transformers are of PCB mounting type. These transformers are designed to operate over the audio frequency range of 20 Hz to 20 KHz. Audio transformers are used in,

- coupling the output of one stage of audio amplifier to the input of the next stage (interstage coupling)
- the amplified audio signal from an amplifier to the speaker of a sound system.

These transformers are said to have *flat frequency response* over the entire audio range. This means that the transformer behaves equally well over the entire range of audio frequencies.

The transformation ratio of audio transformers will be generally less than unity.

These transformers also use a colour coding scheme to identify those used as driver transformers (for inter-stage coupling) or out-put transformers (for amplifier to speaker).

As the colour coding schemes differ from manufacturer to manufacturer, the colour coding is not given in this lesson.

(c) **High frequency transformers:** Refer Fig 6 of Chart 1.

The core of high frequency transformers are made of powdered iron or ferrite or brass or air core (hollow core) as shown in Fig 5 and 7 of this lesson. These transformers are called *Radio frequency transformers* (RFTs) and *Intermediate frequency transformers* (IFTs). These transformers are used for coupling any two stages of high frequency circuits such as radio receivers. The upper frequency limit of these transformers is 30 MHz.

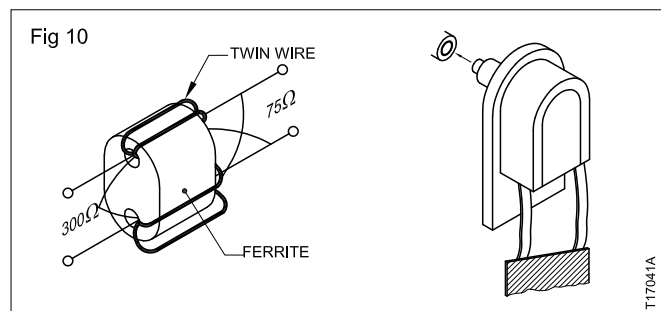
Another speciality of these transformers is that the position of the core can be altered, which results in varied coupling and energy transfer. These transformers also have another electronic component called capacitor connected across the windings in parallel. This results in a different behavior of the transformer at different frequencies. Hence these transformer type are also called *Tuned transformers*.

These transformers are smaller than even audio frequency (AF) transformers. These transformers will generally be shielded/screened using a good conductor (recall lesson on inductors for need of screening).

RFTs and IFTs also have a colour coding scheme to identify their different places of application.

Since the coding schemes vary from manufacturer to manufacturer, the colour coding scheme for RFTs and IFTs are not given in this lesson.

(d) **Very high frequency transformers:** These transformers also have air or ferrite or brass as core material. These transformers are constructed specially to minimize energy losses at very high frequencies. Very high frequency transformers are available in several shapes and designs. Some of these find wide application in Television receivers. Fig 10 illustrates a high frequency transformers used in TV reception.



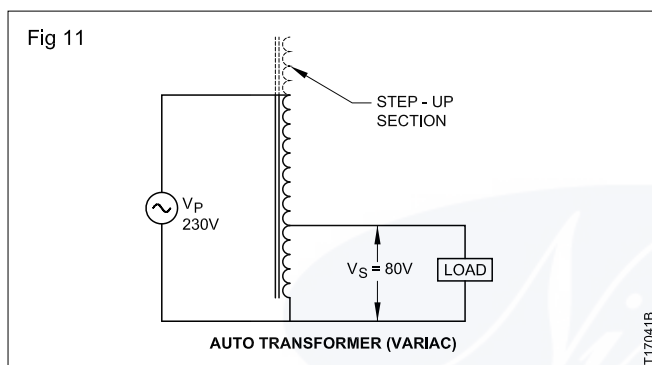
5 Auto-transformers:

Auto-transformer as shown in Fig 11 is a special variety of transformers which have only a single winding. Because of single winding, there is no isolation between primary and secondary side. Auto-transformers are used when isolation between input and output is not important. Auto-transformers can be used for variable voltage operation by using a sliding contact like a potentiometer. But, it is important to note that an auto-transformer does not function as a simple voltage divider.

Auto-transformers are smaller in size and uses less iron than a conventional two winding transformer of the same rating.

Auto-transformers used for variable voltage operation are referred to the trade name of VARIAC.

As shown in Fig 11, auto-transformers has a step-up section (shown in dotted lines) which enables the transformer to develop a variable voltage output from 0 to 270V from a 240V input AC supply.



Auto-transformers are mostly used in laboratories for conducting experiments.

6 Single phase and three phase transformers:

Transformers shown in Fig 4 of Chart are designed for use with single phase AC mains supply. Hence these transformers will have a single primary winding. Such transformers are known as single phase transformers. Transformers are also available for 3 phase AC mains supply. These are known as poly-phase transformers. Refer Fig 7 of Chart 2. In a 3-phase transformer, there will be three primary windings. Three phase transformers are used in electrical distribution and for industrial applications.

7. Classification based on application:

Transformers can also be classified depending upon their application for a specialized work. Since there are innumerable number of application, the types are also innumerable.

However a few of these are listed below:

Current Transformers - used in clip - on current meters, overload trip circuits etc.,

Constant voltage transformers - used to obtain stabilized voltage supply for sensitive equipments

Ignition transformers - used in automobiles

Welding transformers - used in welding equipments

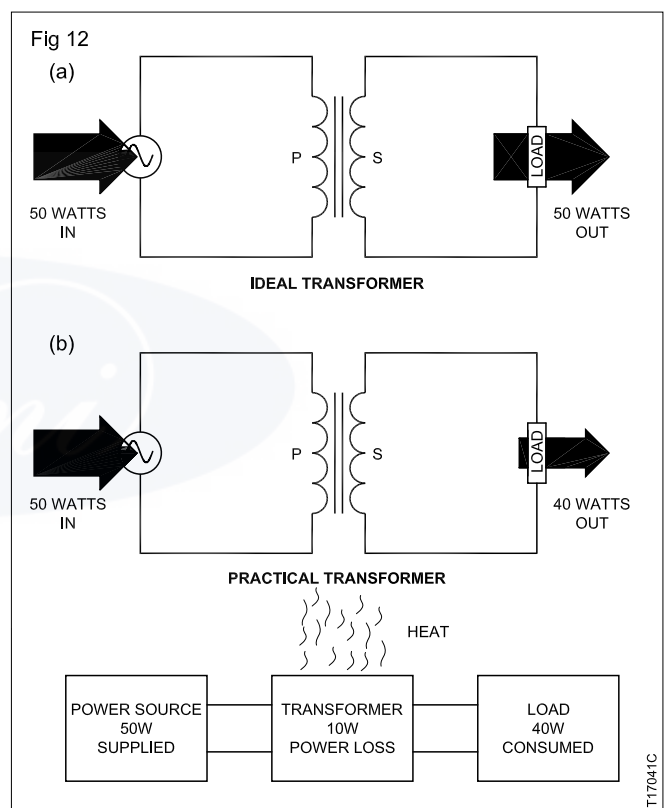
Pulse transformers - used in electronic circuits

Refer reference book for further details on types of transformers and their applications.

EFFICIENCY OF TRANSFORMERS

In an ideal transformer, the power consumed in the secondary is exactly equal to the power consumed in the primary as shown in Fig 12a.

In practice, ideal transformers cannot be made. This is because some amount of power is always wasted in transferring the power from primary to secondary. Hence, the power consumed in primary will always be higher than that available in secondary as shown in Fig 12b. This difference in the power between primary and secondary is lost or wasted as a result of transformer losses.



Transformers can be designed and made so that the transformer losses are minimum. The degree to which any transformer approaches the ideal condition is called the efficiency of the transformer. Efficiency of a transformer is generally expressed in percentage as,

$$\text{Efficiency } h \text{ (in \%)} = \frac{\text{Output power}}{\text{Input power}} \times 100$$

The letter h is spelled as ETA

As an example, efficiency of the ideal transformer shown in Fig 12a is,

$$\text{Efficiency } h \text{ (in \%)} = \frac{\text{Output power}}{\text{Input power}} \times 100$$

$$= \frac{50}{50} \times 100 = 100 \%$$

The efficiency of the transformer shown in Fig 12b is,

$$\begin{aligned} \text{Efficiency } h \text{ (in \%)} &= \frac{\text{Output power}}{\text{Input power}} \times 100 \\ &= \frac{40}{50} \times 100 = 80 \% \end{aligned}$$

From this example, it can be seen that, smaller the output power in relation to the input power, the lower is the efficiency of the transformer.

Efficiency of a transformer is also a measure of the losses in the transformer. The lower the efficiency of the transformer, the higher are the losses in the transformer.

Example: A transformer with 90% efficiency consumes 50 Watts at primary. What is the power delivered at secondary across a purely resistive load.

SOLUTION:

$$\% \eta = \frac{\text{Output power}}{\text{Input power}} \times 100$$

$$\begin{aligned} \text{Output power} &= \frac{\% \eta \times \text{Input power}}{100} \\ &= \frac{90 \times 50}{100} = 45 \text{ Watts.} \end{aligned}$$

Output power = 45 watts.

$$\begin{aligned} \text{Transformer losses} &= \text{Input power} - \text{Output power} \\ &= 50 - 45 = 5 \text{ watts} \end{aligned}$$

In the above example, if the efficiency of the transformer was 95%, then

$$\begin{aligned} \text{Output power} &= \frac{\% \eta \times \text{Input power}}{100} \\ \text{Output power} &= \frac{95 \times 50}{100} = 47.5 \text{ Watts.} \end{aligned}$$

$$\begin{aligned} \text{Transformer losses} &= \text{Input power} - \text{Output power} \\ &= 50 - 47.5 = 2.5 \text{ watts} \end{aligned}$$

LOSSES IN TRANSFORMERS

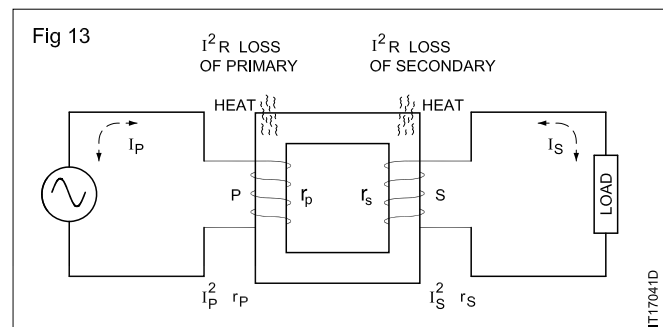
Because of the transformer losses, the transformer heats up when in operation. The losses in the transformer convert some of the electrical energy into heat energy as shown in Fig 12c. As a thumb rule, if a transformer is heating-up while in operation, the losses in the transformer is high.

Most common types of transformer losses which always exist with almost all iron-core transformers are explained below;

1. Copper losses

Transformer windings are made of many turns of copper wire. Copper wire although a very good conductor, still has some resistance (refer Appendix D). The value of this resistance depends upon the type of material and the length of wire. As the number of turns in windings increase, the longer is the length of wire, and greater will be the resistance. As shown in Fig 13, when primary and secondary currents flow through the windings, due to the ohmic resistance of the windings, power (I^2R) is dissipated in the form of heat.

These I^2R losses are called *Copper losses*. Copper losses increase if the currents through primary and secondary increases. Total copper loss in a transformer is equal to;



$$\text{Copper loss} = I_p^2 r_p + I_s^2 r_s$$

Copper losses can be minimised by using a thicker gauge copper wire, but this increases the size, weight and cost of the transformer.

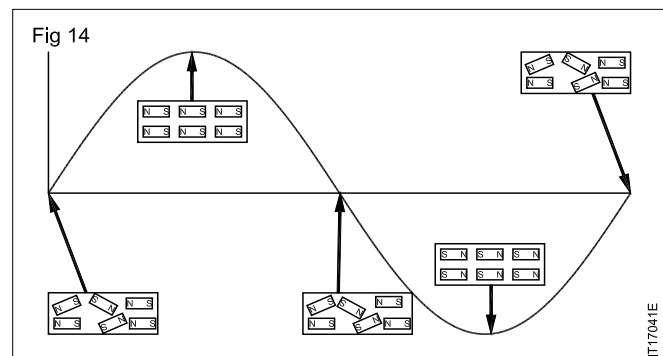
2. Core losses or Iron losses

Core/Iron losses in transformer are due to two different types of losses namely;

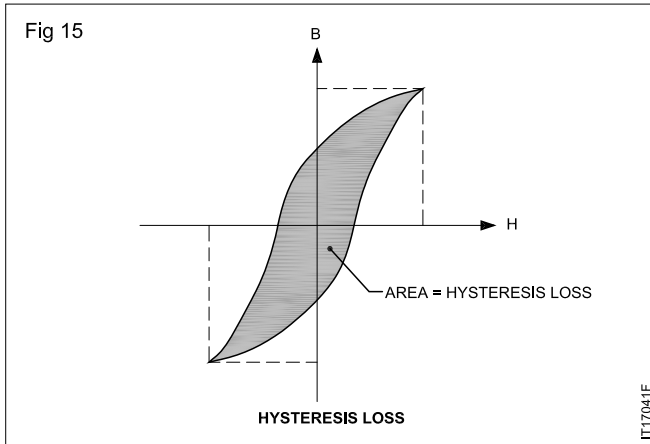
- i) Hysteresis loss
- ii) Eddy current loss

i) Hysteresis loss

The magnetic field in the iron-core of a transformer undergoes a complete reversal 50 times each second for a mains-supply frequency of 50Hz. As shown in Fig 14, every time the polarity of the supply reverses, the molecules of iron with its N-S poles change its direction, such that the direction of magnetic field reverses.



Energy has to be supplied to the molecules of the iron core to make them catch-up with the new direction of magnetic field. This turning around of molecules, or reversing the magnetism of iron core, consumes energy in the form of heat. This loss of energy, appearing in the form of heat, is proportional to the area of the B-H curve or Hysteresis loop of the core material as shown in Fig 15.



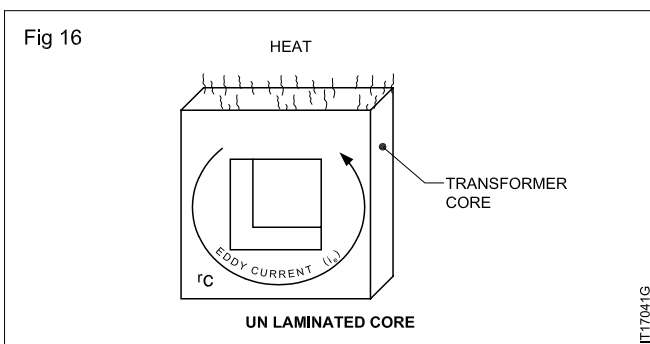
This loss of energy in the primary of the transformer in reversing the magnetism of the iron core is called *hysteresis loss* of the transformer.

It should be noted that air core transformer will not have hysteresis loss as air core transformers do not use magnetic core material.

ii) Eddy current loss

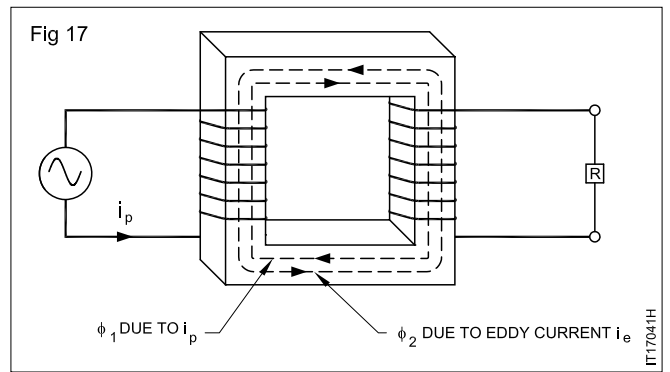
In iron-core transformers, the core material is a conducting material. So, the changing magnetic field of transformer induces a voltage in the core. This induced voltage in the core cause small current to circulate within the core as shown in Fig 16. This current is called eddy current.

The induced eddy current is large if the resistance of the core material is small. Due to this circulating eddy currents and the resistance of the core material, power loss occurs ($P = i_e^2 r_c$) in the form of heat as shown in Fig 16.

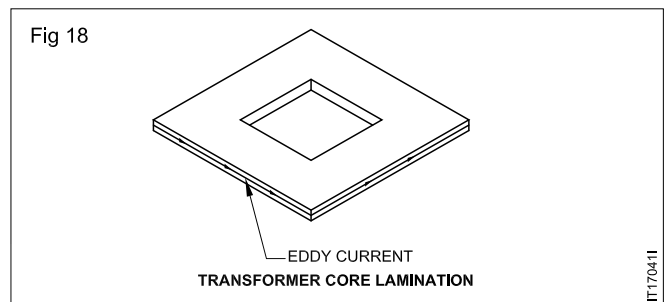


In addition, the induced eddy currents set-up an opposing flux (ϕ_2) in the core as shown in Fig 17. This results in more primary current trying to maintain the magnetic field in the core. This further increases the eddy current and hence the losses due to it.

This loss of power in a transformer due to eddy current in the transformer core is referred to as eddy current loss.

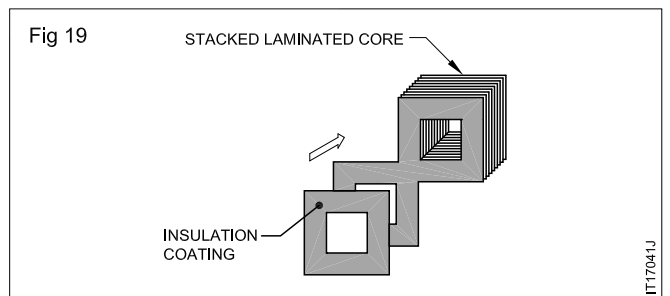


Eddy current loss in a transformer core can be reduced by making the core, into thin flat sections as shown in Fig 18 instead of a solid single piece iron as in Fig 16. These thin flat sections are called laminations.



Since these laminations have very small cross-sectional areas, the resistance offered to the setting up eddy current is greatly increased and hence the loss due to it is also reduced.

Such laminations, are stacked together as shown in Fig 19. These laminations are insulated from each other by means of an insulation coating, generally shellac. Due to the insulation between laminations, the eddy currents can only flow in individual laminations. Hence the overall eddy current loss of the transformer is greatly reduced.



The power loss due to eddy currents is directly proportional to,

- the frequency of current.
- the magnitude of current.

If iron-core transformers are used at high frequencies, the eddy current losses become high. Hence iron-core transformers are not preferred in high frequency applications.

It should be noted that air core transformer will not have any eddy current loss as they do not have core material in which the eddy current can flow.

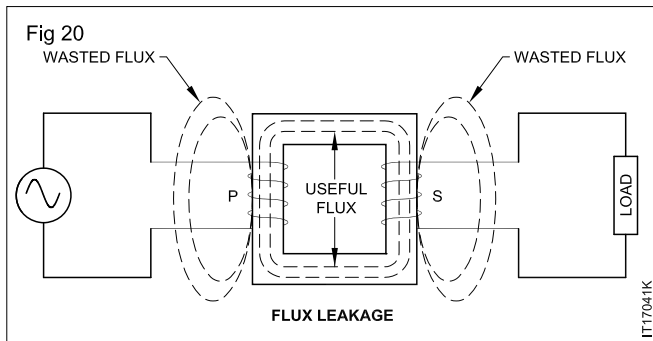
Other losses in transformers

In addition to copper losses and iron losses, transformer have two more types of losses. They are:

- 1) Loss due to flux leakage
- 2) Core saturation loss

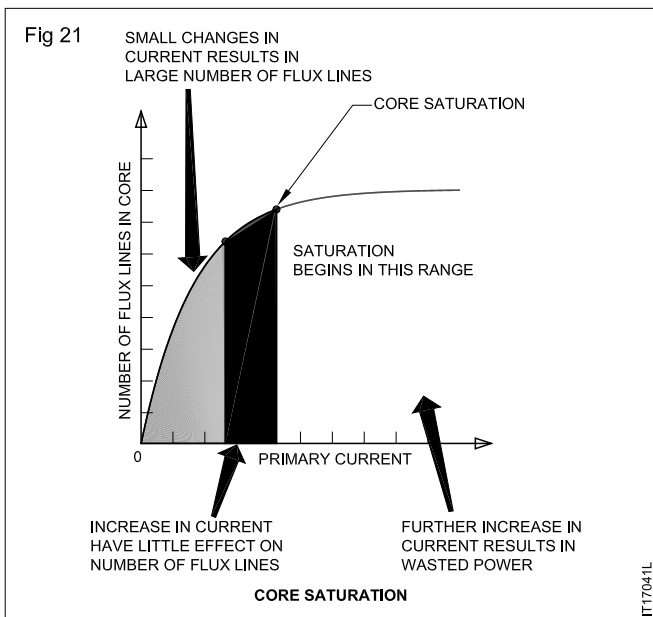
Loss due to flux leakage

As can be seen from Fig 20, in transformers all the flux lines produced by the primary and secondary windings does not travel through the iron core. Some of the magnetic lines leak from the windings and go out into space as shown in Fig 20. These leaked magnetic lines cannot do useful work. This leakage of the flux lines represents wasted energy, reducing the efficiency of the transformer.



Loss due to core saturation

When the current in the primary winding of an iron-core transformer increases, the flux lines generated follow a path through the core to the secondary winding, and back through the core to the primary winding. As the primary current first begins to increase, the number of flux lines in the core increases rapidly shown in light gray shade in Fig 21. Additional increases in primary current will produce only a few additional flux lines less (shown in dark gray shade in Fig 21) than what it should have produced. The core is then said to be saturated. Any further increase in primary current after core saturation, results in wasted power.



Summing the different types of losses in a transformer, the total loss is given by,

$$\text{Total transformer loss} = \text{Copper losses(primary + secondary)} + \text{Iron losses (Hysteresis + eddy current) +}$$

$$\text{Flux leakage loss + Core saturation loss.}$$

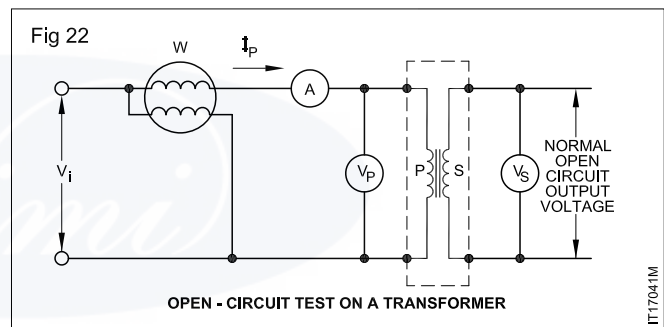
Compared with the other two losses, the flux leakage loss and the core saturation loss are negligible. Also these two losses can be greatly reduced by good transformer design and safe current level operation. Hence, the total losses that occur in a transformer can be found knowing its copper losses and iron losses.

MEASURING TRANSFORMER LOSSES

To determine losses in a transformer, its turns ratio and efficiency, two simple tests are conducted. These tests are, the *no-load test* and the *full-load test*.

No-load test or open circuit test (O-C test)

Fig 22 shows the circuit arrangement for O-C test on a transformer.



The ac input voltage (V_i) is set at a rated primary voltage. The input power (P_i) is measured by the wattmeter (W). The input current (I_p) is measured by ammeter.

The open-circuit secondary voltage (V_s) is measured by voltmeter.

Since the secondary is open there is no current in secondary.

As the transformer secondary is open-circuited ($I_s=0$), the primary current (I_p) is very small. Since I_p is very small, the voltage drops across the ammeter and wattmeter can be neglected. So the input voltage (V_i) can be taken as primary voltage (V_p). Therefore, the ratio of the two voltmeter readings gives turns ratio of the transformer.

$$\text{Turns ratio of transformer} = \frac{N_p}{N_s} = \frac{V_p}{V_s}$$

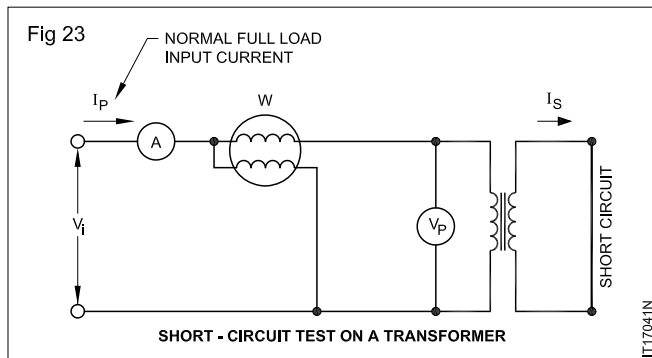
The input power (P_i) measured by the wattmeter (W) gives the total transformer *core losses* because with very small primary current and zero secondary current the copper losses in the windings are negligible and hence can be taken as zero.

$$\begin{aligned}\text{Total losses in a transformer} &= \text{Copper loss} + \text{Iron loss} \\ &= 0 + \text{Iron loss}\end{aligned}$$

With copper loss being zero, the input power measured on the wattmeter (W) is the total transformer Core losses or Iron losses (W_i).

Full load test or short circuit test (S-C test)

Fig 23 shows the circuit arrangement for S-C test on a given transformer.



With the secondary terminals shorted, the input voltage (V_i) is increased slowly from zero till the ammeter in the primary circuit indicates rated full-load primary current, I_p . When this occurs, the rated full load secondary current I_s will be circulating in the secondary winding.

Because the secondary terminals are shorted, the voltage required at primary, V_p to produce full-load primary and secondary current is just around 3% of the rated input voltage (V_i).

In this condition, the wattmeter measuring input power (P_i) indicates the full-load copper losses for the reasons given below;

- With a low level of input voltage (3% of rated), core flux is minimum. Hence the core losses are so small that they can be neglected and taken as zero.
- Since the winding, both primary and secondary are carrying rated full-load currents, the input is supplying the rated full-load copper losses only.

$$\text{Total losses} = \text{Copper loss} + \text{Iron loss}$$

$$\text{Total losses} = \text{Copper loss} + 0$$

With Iron loss being zero, the input power measured (W_c) on the wattmeter is the total transformer copper loss at rated full-load current.

Using the results of the SC test, the phase angle difference (ϕ) between the current and the voltage can be determined as given below;

$$\text{Power factor, } \cos \phi = \frac{\text{True power}}{\text{Apparent power}}$$

True power is the power read on the wattmeter, apparent power is given by, $V_p \times I_p$ voltamps.

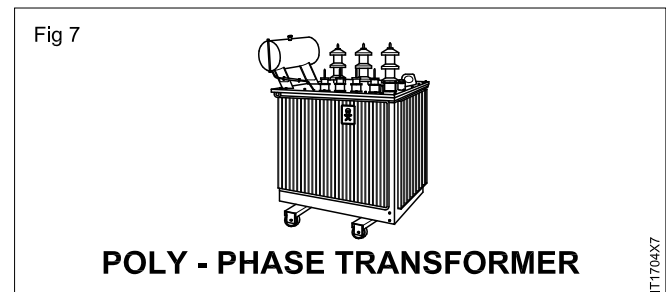
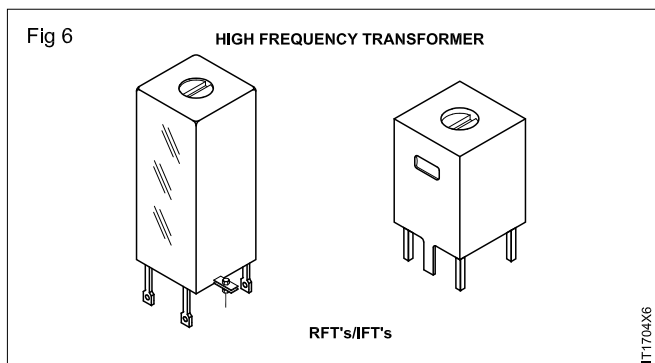
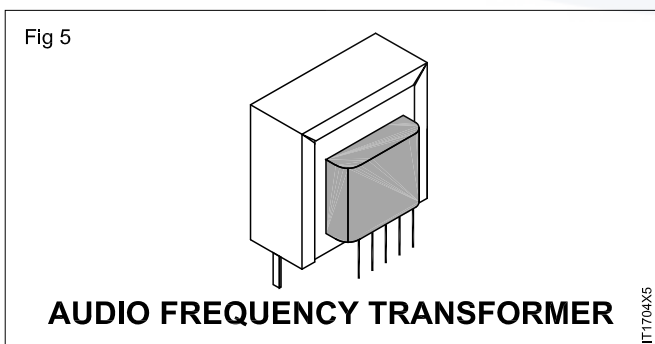
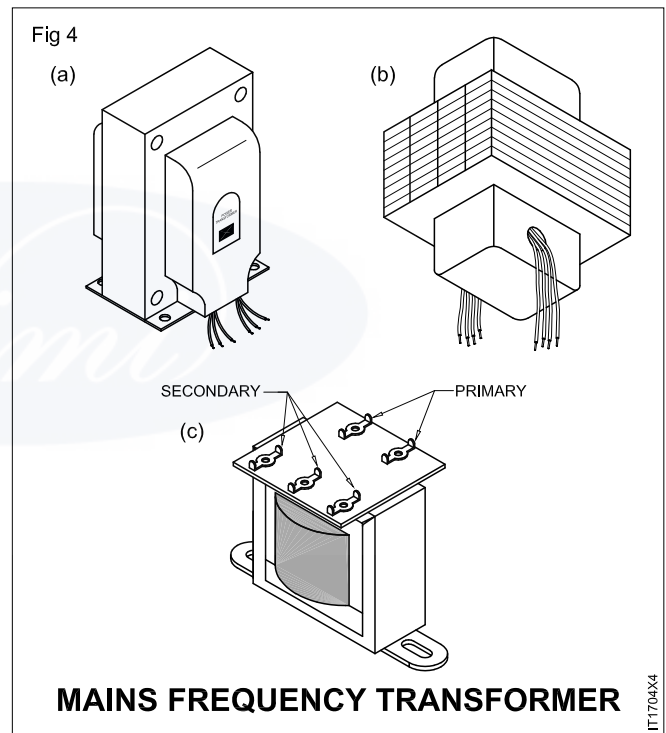
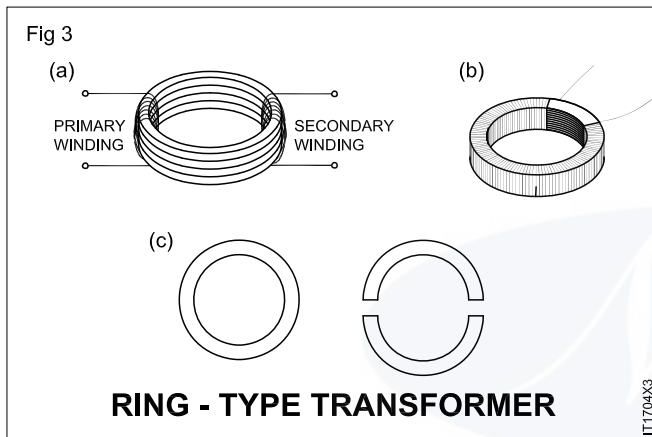
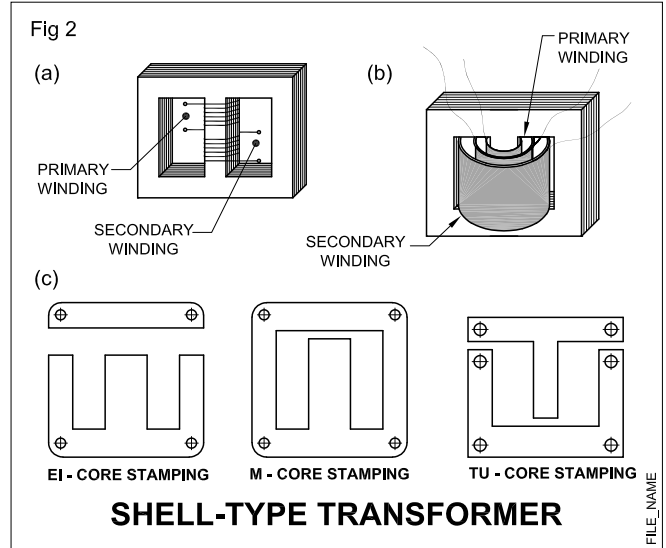
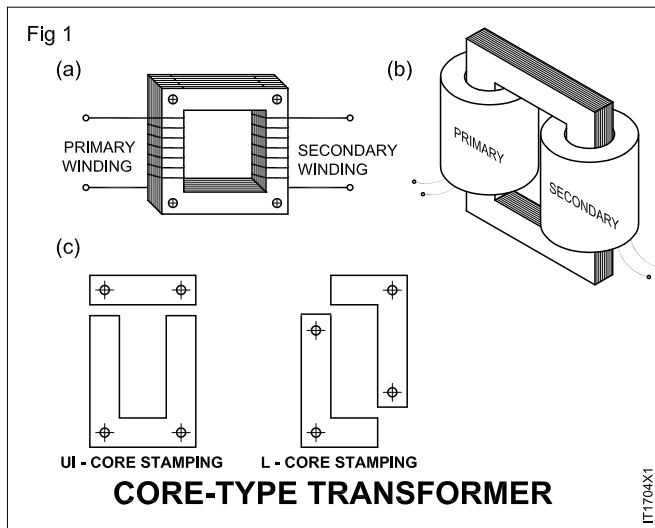
Therefore, in SC test

$$\cos \phi = \frac{W_c}{V_p \cdot I_p}$$

$$\phi = \cos^{-1} \left(\frac{W_c}{V_p \cdot I_p} \right)$$

CHART- 1

TYPES OF TRANSFORMERS



Voltage regulation, efficiency and phase relationship in transformers

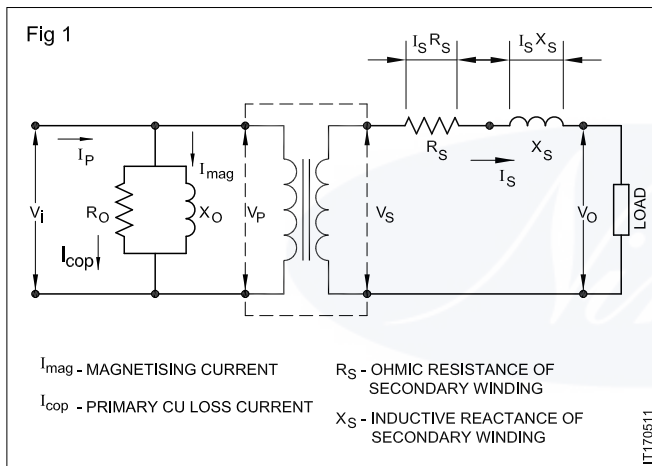
Objectives: At the end of this lesson you shall be able to

- explain the meaning of regulation in transformers
- determine regulation from OC and SC test data
- explain the method to determine the efficiency of a transformer from OC and SC tests
- describe the phase relationship between primary and secondary with reasons.

Voltage Regulation

For simplicity in understanding, in all the previous lessons on transformers, the resistive and inductive effect of the primary and secondary windings were neglected. Also, the effect of load current on the magnitude secondary voltage was neglected. In a practical situation, the secondary voltage of a transformer decreases as the load/load-current is increased.

Consider the equivalent circuit of a transformer shown in Fig 1.



From the circuit in Fig 1, the secondary current I_s produces voltage drops $I_s R_s$ and $I_s X_s$ across the resistive and reactive components of the secondary winding. Consequently, the output voltage V_o is less than V_s .

When load is not connected at the output terminals of the transformer, no secondary current flows, and hence, no voltage drops occur across R_s and X_s . Hence, V_o equals V_s . Thus, the transformer secondary voltage is greatest on no-load.

Under loaded condition, the voltage drops across the resistive and reactive components of the secondary winding, reducing V_o . The larger the load current, larger will be the drop across the resistive and reactive components of the secondary and hence, smaller will be the value of V_o .

The percentage change in output voltage V_o , from no-load to full load is termed the voltage regulation of the transformer.

% Voltage regulation =

$$\frac{V_o(\text{No - load}) - V_o(\text{Full - load})}{V_o(\text{Full - load})} \times 100\%$$

Ideally, there should be no change in V_o from no-load to full-load, (i.e., regulation = 0%). For best possible performance, the transformer should have the lowest possible percentage regulation.

In some text books, the regulation discussed above is termed as “% Regulation-Up” some books also use, the term “% Regulation-Down” given by,

$$\% \text{ voltage regulation down} = \frac{V_{o(NL)} - V_{o(FL)}}{V_{o(NL)}}$$

For example, if a transformer has an output of 13 V when on no-load and has an output of 11.8 V when on its rated resistive load, the regulation of the transformer is,

$$\% \text{ Voltage regulation} = \frac{V_{o(NL)} - V_{o(FL)}}{V_{o(FL)}} \times 100\%$$

$$= \frac{13 - 11.8}{11.8} \times 100\% = 10\%$$

Finding regulation from OC and SC test results

$$\text{Voltage regulation} = \frac{V_{o(NL)} - V_{o(FL)}}{V_{o(FL)}} \times 100\%$$

$V_{o(NL)}$ is the secondary voltage in test obtained from OC test.

$V_{o(FL)}$ is the secondary voltage under rated full load.

$V_{o(FL)}$ can be calculated knowing turns ratios and the data obtained in SC test as follows;

$$V_{o(FL)} = V_{o(FL)} \times \frac{N_s}{N_p}$$

$$\text{where, } V_{p(FL)} = V_{p(NL)} - V_{i(SC)}$$

where, $V_{i(SC)}$ is the voltage applied to primary at full load in SC test.

Calculating efficiency from the O-C, S-C test data

$$\text{Efficiency } \eta \% = \frac{\text{Output power } (P_o)}{\text{Input power } (P_i)} \times 100$$

Since, Input power (P_i) = Output power (P_o) + Losses

$$\eta \% = \frac{P_o}{P_o + \text{losses}} \times 100 \quad \dots\dots\dots[1]$$

Efficiency η of a transformer can also be found as follows

$$\eta = \frac{\text{Output power } (P_o)}{\text{Input power } (P_i)}$$

Since, Output power (P_o) = Input power - Losses

$$\eta = \frac{\text{Input power} - \text{Losses}}{\text{Input power}}$$

$$\eta = 1 - \frac{\text{Losses}}{P_i}$$

Output power P_o is given by,

$$P_o = V_{s(FL)} \times I_{s(FL)} \times \cos\phi_s$$

$I_{s(FL)}$ is the secondary current under full load condition. This can be calculated using the data obtained in S-C test as follows,

$$I_{s(FL)} = I_{p(FL)} \times \frac{N_p}{N_s}$$

For simplicity, assuming the effect of inductance in secondary is negligible $\cos\phi_s$ can be taken as 1.

$$\text{therefore, } P_o = V_{o(FL)} \times I_{s(FL)} \quad \dots\dots\dots[2]$$

Core losses is given by wattmeter(W_i) reading in O-C test

$$\dots\dots\dots[3]$$

Full load copper losses is given by wattmeter reading in SC test

$$\dots\dots\dots[4]$$

Knowing [2], [3] and [4], the efficiency of the transformer can be calculated using ..[1].

Example: A 36VA rated step down transformer (230V:12V, 3A secondary) on OC and SC tests gave the following results;

OC test: $V_p = 220V$, $I_p = 0.068A$, $W = 2$ watts

SC test: $V_p = 21V$, $I_p = 0.3A$, $W = 6$ watts

Find the efficiency of the transformer at the rated secondary load of 36VA. Assume secondary load P.F($\cos\phi$) = 1.

From OC test:

$$\text{Apparent power} = V_p \cdot I_p = 220 \times 0.068 = 14.96 \text{ VA}$$

$$\text{True power} = W = 2W$$

$$\cos \phi = \frac{\text{True power}}{\text{Apparent power}} = \frac{2}{14.96} = 0.1337$$

$$\cos\phi = 0.1337$$

$$\phi \simeq 82^\circ$$

Iron losses = Wattmeter reading in O.C test = 2 watts

From SC test:

Short circuit voltage $V_{s(FL)}$ is given by;

$$V_{s(FL)} = \frac{V_{p(SC - \text{test})}}{V_{p(SC - \text{test})}} = 0.0954 \text{ volts}$$

$$I_{s(FL)} = \frac{V_{p(OC - \text{test})} \times I_{p(SC - \text{test})}}{V_{p(SC - \text{test})}} = \frac{(220V)(0.3A)}{21V} = 3.14$$

Copper losses = Wattmeter reading in SC-test = 6 watts

Total transformer losses = Copper loss + Iron loss

$$= 6 \text{ watts} + 2 \text{ watts}$$

$$= 8 \text{ watts.}$$

Efficiency of the transformer at rated load of 36VA is

$$\eta \% = \frac{36}{36 + 8} \times 100\% = 81.82\%$$

These calculations have not taken into account the effect of inductive reactance of the primary and secondary circuit. If the load connected to the secondary of a transformer is going to be a highly inductive load, then it is necessary to consider the power factor ($\cos\phi$). For details refer reference books listed at the end of this book for exact calculations of efficiency of transformers from the data obtained by O-C and S-C tests.

PHASE RELATIONSHIP BETWEEN PRIMARY AND SECONDARY

With Open secondary winding

This means,

- the primary current lags behind the applied voltage V_i by 90° as shown in Fig 2b.
- From Lenz's law the back-emf produced in the primary, which opposes the cause, therefore lags behind the primary current by 90° as shown in Fig 2c.

The voltage induced in secondary is maximum when the primary back-emf is maximum. That means,

- the secondary voltage lags behind the primary current by 90 degrees and hence the secondary voltage (V_s) is 180° out of phase with the primary voltage.

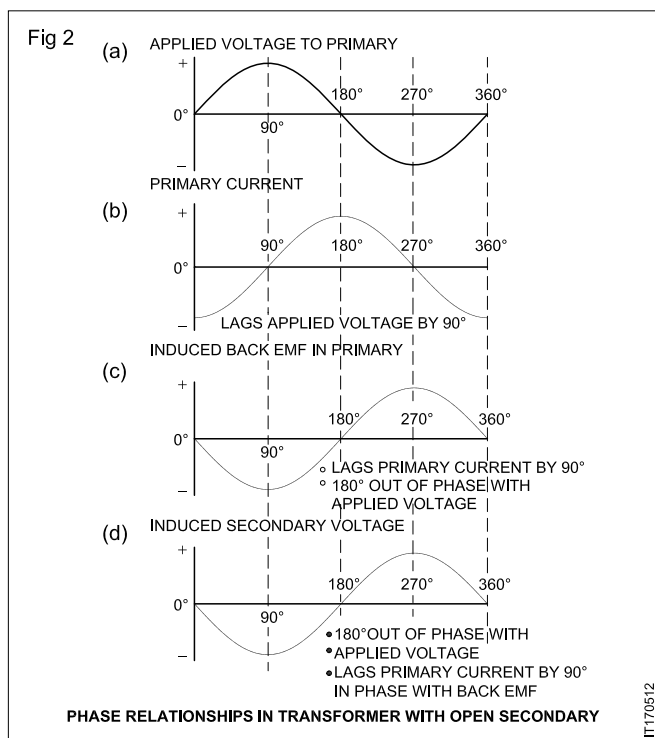
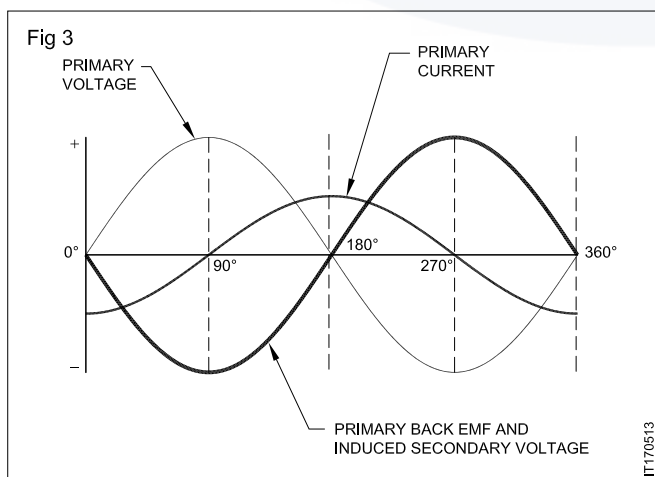


Fig 3 shows a combined illustration of the phase relationship between primary and secondary.

With loaded secondary

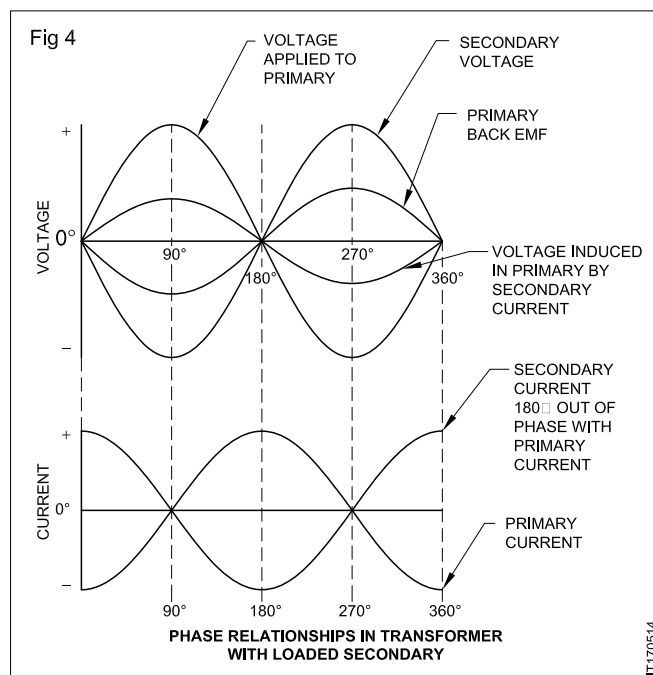
When a load is connected to the secondary of a transformer, current flows in the secondary. As in any inductance,

- the current through the secondary winding lags behind the secondary voltage that produces it by 90° degrees.



Since the secondary voltage lags behind the primary current by 90° (Fig 4), and since the secondary current lags behind the secondary voltage by 90° ,

- the secondary current is 180° out of phase with primary current as shown in Fig 4.



As the secondary current changes, it generates its own magnetic field, whose flux lines oppose those of the magnetic field created by the primary current. This reduces the strength of the primary magnetic field. As a result, less back-emf is generated in the primary. With less back-emf to oppose the applied voltage, the primary current increases. The amount of increase in primary current is directly proportional to the amount of increase in secondary current. Thus, **when secondary current in a transformer increases, the primary current automatically also increases.** And when secondary current decreases, the primary current also decreases.

Applying rated primary voltage, if the secondary of a transformer is shorted, excessive current will flow in the primary as well as in the secondary. This excessive current will not only burn out the transformer, but there is a possibility that the source supplying power to the primary would also be damaged.

TRANSFORMER RATINGS

1 Voltage rating

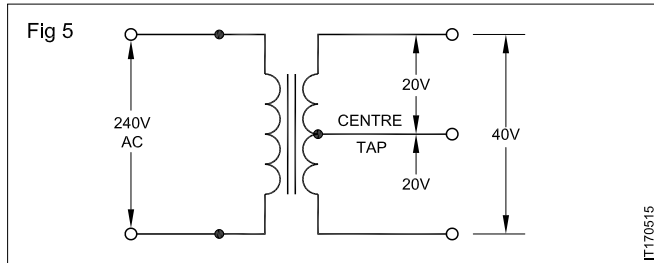
Transformer manufacturers always specify the voltage rating of both the primary and secondary windings.

Operating the primary above rated voltage usually causes the transformer to overheat. More serious is the additional stress placed on the transformer insulation by the higher primary and secondary voltages. Operating the primary below rated voltage does no harm, but this makes the secondary voltages lower than rated values.

The rated voltages of the secondaries are specified for full-load conditions with rated primary voltage. With no-load, the secondary voltage is slightly higher than rated voltage (usually 5 to 10 percent higher).

Manufacturers indicate the specification of centre-tapped secondaries in different ways. For example, the secondary in Fig 5 may be specified in any of the following ways:

- 1) 40 V Centre tapped
- or 2) 20 V - 0 - 20 V
- or 3) 20 V each side of center



2 Current rating

Manufacturers usually specify current ratings for secondary windings only. This is because, as long as the secondary current rating is not exceeded, the primary current-carrying capacity cannot be exceeded.

Exceeding the current rating of a secondary causes its voltage to fall below rated value. But more serious than this slight decrease in secondary voltage is, the increase in, I^2R (copper-loss) loss in the secondary. The increased I^2R loss causes the winding to overheat and eventually destroys the transformer. Increase in copper-loss reduces the efficiency of the transformer.

3 Power rating (W rating)

Some manufacturers specify a power rating (in watts) for their transformers. This is understood to be the power the transformer can deliver to a resistive load. Thus the power rating is merely the product of the current rating and the voltage rating of the secondary ($P=VI$). For multiple-secondary transformers, the power rating is the sum of the powers available from the individual secondaries ($P_T = P_1 + P_2 + \text{etc.}$).

The total power(P_T) cannot be taken from a single secondary on a multiple-secondary transformer. The current rating of the individual secondaries must not be exceeded.

4 Voltampere rating (VA rating)

The Voltampere rating(VA) of a transformer is an apparent power rating. It is applicable to any type of load Resistive, Reactive, or its combination(impedance). The VA rating, like the power rating, is given for the total transformer secondary instead of individual secondary windings. In a multiple-secondary transformer, the total voltampere rating cannot be taken from a single secondary.

From the details given above, it can be seen that, manufacturers use many ways to rate transformers. From users point, you will never overload a transformer or exceed any of its rating if you observe the following two basic rules,

- 1 Never apply more than the rated voltage to the primary winding.
- 2 Never draw more than the rated current from secondary winding(s).

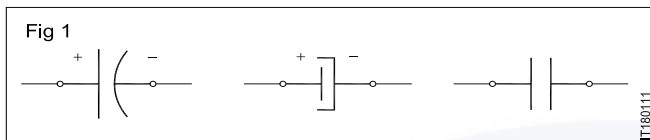
For ease of understanding the phase relation ship between voltages and currents in primary and secondary of a transformer, consider a transformer having an open secondary. Referring to Fig 2, with open secondary, the primary winding works similar to that of a inductor.

Capacitors - construction & testing

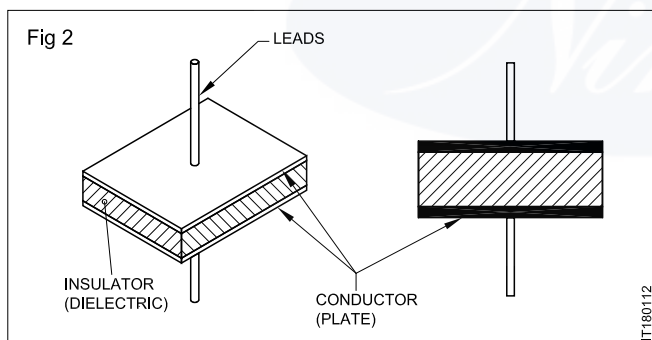
Objectives: At the end of this lesson you shall be able to

- state the basic function of capacitors and meaning of capacitance
- describe how capacitor stores energy and define the unit
- list factors that determine and state the effect of A, d, k on capacitance value
- list functions of dielectric in a capacitor
- list main types of fixed value capacitors and minimum specifications
- state the important of DCWV in capacitors
- list common defects and their causes in capacitors
- describe the term *capacitor action*
- explain the two common methods of testing capacitors.

Capacitors are electronic components which can store electric energy in the form of electric charge. The charge storage ability of a capacitor is called the **Capacitance** of a capacitor. Symbols used to represent capacitors are shown in Fig 1. Alphabet 'C' is used to represent the capacitance of a capacitor.

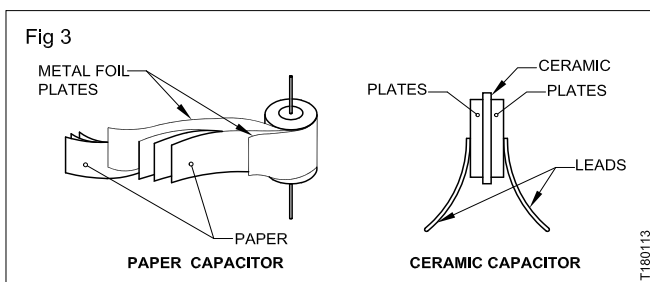


A simple capacitor consists of two pieces of conductors separated by an insulator as shown in Fig 2.



In capacitors the conductors shown in Fig 2 are called **plates** and the insulator is called **dielectric**.

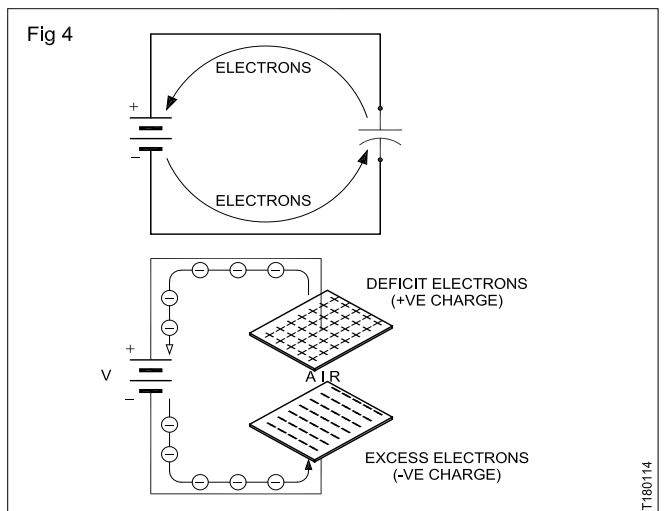
The plates of a capacitor can be of any size and shape and the dielectric may be any one of several insulator materials. Depending on the type of insulator/dielectric used capacitors are called as paper, mica, ceramic, glass, polyester, air electrolyte capacitors etc., as shown in Fig 3.



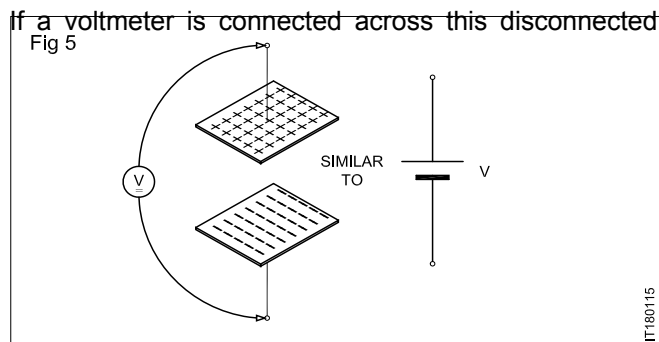
Capacitor action of storing charge

When electric charge is forced on to the plates of a capacitor by some energy source, such as a battery, the capacitor stores these charges.

When a capacitor is connected to a battery as shown in Fig 4, electrons from the negative terminal of battery move through the connecting leads and pile up on one of the plates of the capacitor. At the same time free electrons from the other plate of the capacitor (remember that plates of a capacitor are conductors having free electrons) move through the connecting lead to the positive terminal of the battery. This process is known as 'charging of capacitor'. As the process of charging continues, the net result is that, one plate of the capacitor ends up with excess of electrons (Negative charge) and the other plate with deficiency of electrons (Positive charge). These charges on the plates of the capacitor represent a voltage source similar to that of the charges on the terminals of a battery/cell. The process of charging stops once the energy stored on the capacitor develops a voltage equal to that of the battery.



It is important to note that during the process of charging, although electrons were moving from and to the capacitor plates causing current flow in the circuit (you can connect an ammeter to measure it), no electrons moved nor did current flow from one plate through the dielectric to the other plate of the capacitor. The charging current through the circuit stops when the voltage across the capacitor becomes equal to, and in opposition to, the battery voltage. This charged capacitor can be disconnected from the circuit and used as a new energy source as shown in Fig 5.



charged capacitor, the voltmeter reads the voltage equal to that of the battery which charged it.

If a lamp is connected across this charged capacitor, the bulb glows for a moment indicating current flow through it.

[H.I: The instructor to demonstrate charging of a capacitor, voltage across a disconnected charged capacitor and discharge of a charged capacitor through a lamp using a suitable demonstration circuit.]

The charge stored in the capacitor is sufficient to supply current through the bulb only for a short duration after which the charge filed up on the capacitor plates gets exhausted. A capacitor has limited use as a primary storage device of energy for two reasons:

- 1 For its weight and size, the amount of energy it can store is very small when compared with that of a battery.
- 2 The voltage available from the capacitor diminishes rapidly as energy is removed from the capacitor.

Unit of capacitance

The ability of capacitor to store electrical energy in the form of electrostatic field is known *capacitance*. The unit used to measure capacitance is **Farad** abbreviated as **F**.

A capacitor is said to have a capacitance(C) of 1 Farad, if it stores a charge(Q) of 1 coulomb when a voltage(V) of 1V is applied across its plates.

Therefore, capacitance can be mathematically expressed as,

$$\text{Capacitance} = \frac{\text{Charge}}{\text{Voltage}}$$

$$C = \frac{Q}{V} \text{ Farads}$$

Farad(F) is a very large quantity of capacitance. As most circuits use capacitance values much lower than one farad (F), smaller quantities of capacitance given below are generally used:

$$1 \text{ Microfarad or } 1\mu\text{F} = 1/1000000 \text{ F or } 10^{-6} \text{ farads}$$

$$1 \text{ Nanofarad or } 1 \text{ nF} = 1/10^9 \text{ F or } 10^{-9} \text{ farads}$$

$$1 \text{ Picofarad or } 1\text{pF} = 1/10^{12} \text{ F or } 10^{-12} \text{ farads}$$

Example: What is the capacitance (C) of a capacitor that requires a charge (Q) of 0.5 coulombs to build a voltage(V) of 25 volts across its plates?

SOLUTION

Given: Charge (Q) = 0.5 Coloumb

Voltage (V) = 25 Volts

Using the formula,

$$\text{Capacitance, } C = \frac{Q \text{ Coloumbs}}{V \text{ Volts}} \text{ Farads}$$

$$\text{Capacitance, } C = \frac{0.5}{25} = 0.02 \text{ Farads}$$

Factors that determine the value of capacitance

The capacitance of a capacitor is determined by the following three main factors;

- 1 Area of the plates
- 2 Distance between the plates
- 3 Type of dielectric material (dielectric constant k)

Chart 1 at the end of this lesson illustrates the relationship between the above factors and the value of capacitance.

In addition to the above factors affecting the value of capacitance, the temperature of the capacitor also affects the capacitance although not very significantly. Increase or decrease in temperature affects the characteristics of dielectric material which in-turn increases or decreases the capacitance value. Some dielectrics cause an increase in capacitance as temperature increases. These are called positive temperature coefficients, abbreviated as P. Other dielectric materials have negative temperature coefficient, abbreviated as N, in which case, increase in temperature decreases the capacitance. There are dielectric materials having zero temperature coefficient abbreviated as NPO. The temperature coefficient of a capacitor is specified by the capacitor manufacturer in parts per million per degree Celsius (PPM).

The following expression gives the relation between the three factors that determine the value of capacitance of a capacitor;

$$C = \epsilon_r \epsilon_0 \frac{A}{d} \text{ Farads}$$

The term ϵ_0 is the permittivity of free space (air) = $8.85 \times 10^{-12} \text{ C}^2/\text{Nm}^2$ and ϵ_r is called the relative permittivity of the dielectric material.

The expression for capacitance (C) of a capacitor can also be written as,

$$C = k \epsilon_0 \frac{A}{d} \text{ Farads}$$

The ratio of the capacitance with dielectric to the capacitance with air is called relative permittivity or dielectric constant, k.

Substituting the value of ϵ_0 in the above equation, value of a capacitor using any dielectric can be found using the formula;

$$C = (8.85 \times 10^{-12}) k \frac{A}{d} \text{ Farads}$$

where,

C = Capacitance in farads

$(8.85 \times 10^{-12}) = \epsilon_0$ (permittivity of air)

k = dielectric constant of the insulator used between the plates

A = area of one side of the plate in square meters, m^2

d = distance between the plates in meters, m

Example: Two metal plates, each 5 x 6 cms are separated from each other by 1mm. Calculate the capacitance if the dielectric material used between the plates was,

- 1) air
- 2) glass

SOLUTION:

- 1 From table in Appendix 'D'

$$k_{\text{air}} = 1$$

$$C = (8.85 \times 10^{-12}) k \frac{A}{d}$$

$$= (8.85 \times 10^{-12}) \times 1 \times \left(\frac{5 \times 10^{-12} \text{ m} \times 6 \times 10^{-12} \text{ m}}{1 \times 10^{-3} \text{ m}} \right)$$

$$= (8.85 \times 10^{-12}) \times 1 \times$$

$$= 26.55 \times 10^{-12} \text{ Farads}$$

$$= 26.55 \text{ pico farads}$$

$$\mathbf{C = 26.55 \text{ pF}}$$

- 2 From table in Appendix 'D'

$$k_{\text{Glass}} = 5$$

$$C = (8.85 \times 10^{-12}) \times 5 \times \left(\frac{5 \times 10^{-12} \text{ m} \times 6 \times 10^{-12} \text{ m}}{1 \times 10^{-3} \text{ m}} \right)$$

$$= 5 \times 26.55 \text{ pF}$$

$$\mathbf{C = 132.75 \text{ pF}}$$

Dielectric constants, k of commonly used insulators are given in Appendix D.

Working voltage or voltage rating of capacitor

The dielectric strength of the insulating material used between the plates of a capacitor gives the capacitor the ability to withstand a potential difference between the plates without causing arcing. Therefore, a specific capacitor using a specific type of dielectric can withstand only up to a specific voltage across it. If the voltage is further increased, the dielectric breaks down or gets punctured. This causes a burn out or a hole in the dielectric material permanently damaging the capacitor.

This maximum voltage that a capacitor can withstand is listed as one of the specifications of capacitors as *direct current working voltage*, DCWV. As an example: if a capacitor has a DCWV of 100 volts, it can be operated at 100 volts for long periods of time without any deterioration in the working of the capacitor. If the capacitor is subjected to 125V or 150V dc, the dielectric may not break down immediately but the life of the capacitor gets greatly reduced and may become permanently defective any time.

Function of a dielectric in a capacitor

- 1 Solves the mechanical problem of keeping two metal plates separated by a very small distance.
- 2 Increases the maximum voltage that can be applied before causing a breakdown, compared with air as dielectric.
- 3 Increases the amount of capacitance, compared with air, for a given dimension of plates and the distance between them.

Types of capacitors

Capacitors can be classified under two main categories:

1 Fixed value capacitors

The capacitance value of these capacitors is fixed at the time of manufacture. This value cannot be varied/altered by the user.

2 Variable capacitors

The capacitance of such capacitors can be varied between the specified minimum to the specified maximum values by the user.

Amongst fixed value capacitors, many different types of capacitors are manufactured to satisfy the needs of the electronic industry. These different types of capacitors are named according to the

- 1 Type of dielectric material used in capacitor

Example:

- a) If paper is used as dielectric, the capacitors are called paper capacitors.
- b) If ceramic is used as dielectric, the capacitors are called Ceramic capacitors.

- 2 Type of construction of the capacitor

Example:

- If the foils of the conductor and dielectric are rolled to form a capacitor, such capacitors are called *Rolled foil capacitors*.
- If the plates and dielectric are in the form of Discs, such capacitors are called *Disc capacitors*.

Different types of fixed value capacitors, their sub types, available values, rated voltage and a few applications are given in Chart 2 at the end of this lesson. Also refer to Chart 3 for illustration of some of the popular fixed value capacitors.

Specifications of capacitors

While ordering capacitors, one has to indicate the specifications needed to ensure that the desired capacitor is received. The minimum specifications to be indicated while purchasing/ordering capacitors for general use are;

1 Type of capacitor

For example: Ceramic, disc, styroflex, electrolytic and so...on.

2 Capacitance value

For example: 100 μ F, 0.01 μ F, 10pf and so....on.

3 DC working voltage rating (DCWV)

For example: 100 μ F-12V, 100 μ F-100V, 0.01 μ F-400V and so...on.

4 Tolerance

Like resistors, capacitors also have tolerances over its rated value. Tolerance of capacitors may range from $\pm 1\%$ to $\pm 20\%$. Some capacitors may have tolerance specified as -20% , $+80\%$.

5 Temperature coefficient

The temperature coefficient of capacitors can be Positive (P), Negative (N) or zero (NPO). The temperature coefficient is expressed in parts per million (ppm) per degree Celsius.
Example: +750ppm

The above listed capacitor specifications, the manufacturer's name/code and the month and year of manufactured are generally marked on all capacitors.

While identifying the specification of a given capacitor the following points are to be noted.

- The type of capacitor may not be marked on the capacitors. In such cases the type has to be identified depending on the physical appearance of the capacitor.
- Capacitance value of capacitors may be marked directly as μ F, pF etc or might be colour coded. It should be noted that the capacitance value formed using colour code is always in Picofarads (PF).

Different capacitor manufacturers adopt different coding schemes for indicating the capacitance value on the capacitor. Refer Appendix D for different colour coding schemes.

- The DC working voltage (DCWV) will be marked directly as,
16V - meaning 16V DC - marking scheme used in electrolytic capacitors
400V - meaning 400V DC - marking scheme used in paper capacitor
2 - meaning 200V DC - marking scheme used in ceramic disc/plastic capacitors

Refer Appendix D for further details.

- The tolerance of capacitors will be marked either directly as 10%, 20% or in a coded form as K (meaning 10%), M-meaning 20% and so...on. Refer to PTB for other tolerance codes.
- The temperature coefficient will generally be marked in the coded form as, P, N or NPO.

Many capacitors may not have all the markings on their body. However, specification can be obtained for a particular type from its manufacturer.

CONSTRUCTION OF CAPACITORS

Although all capacitors in general have plates and dielectrics, depending upon the type of dielectric used and the type and nature of the plates, construction of capacitors varies drastically. Some of the main types of capacitors and their construction are illustrated in Chart-4 at the end of this lesson.

For further details on the construction of capacitors refer reference books listed at the end of this book.

COMMON DEFECTS IN CAPACITORS

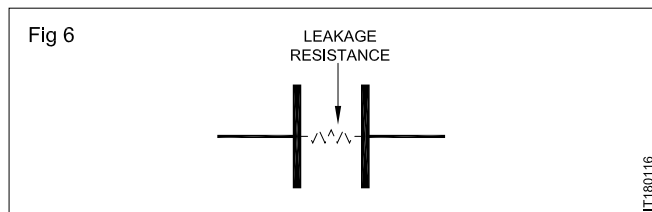
Short circuited capacitors: In the course of normal usage, capacitors can become short-circuited/shorted. This is because of the deterioration of the dielectric used due to aging. Usually such a defect occurs when the capacitor is used over a period of years under the stress of changing voltage across it. This period gets reduced if the capacitor is operated at higher temperatures. Short-circuiting of capacitors is more common in paper and electrolytic capacitors than in the other types. Short circuiting of capacitor also occurs due to puncturing of the dielectric when voltages much higher than its rated are applied. A shorted capacitor cannot store energy.

Open capacitors: A capacitor may become open due to loose/broken lead connections or due to the electrolyte. In electrolytic capacitors, the electrolyte develops high resistance with age, particularly when operated at high temperatures. After a few years of usage, the electrolyte may dry up resulting in open-circuit of the capacitor. An open capacitor cannot store energy.

The storage (shelf) life of wet type electrolytic capacitors is small because the electrolyte dries up over period of time.

Leaky capacitors/leakage resistance

Theoretically, the current that flows in a pure capacitive circuit results from the charge and discharge currents of the capacitor. The dielectric, which is an insulator, should prevent any current flow between the plates. However, even the best dielectric conduct very small current. The dielectric, then, has some high value of resistance, known as leakage resistance. This leakage resistance, as shown in Fig 6, allows some leakage current to flow. This leakage current tends to reduce the capacitance value.



In a good capacitor, the leakage resistance is generally of the order of several tens of megohms, and hence can be considered negligible for most applications. As the capacitor ages, the leakage resistance could reduce. Generally, the leakage resistance is lower with high-value capacitors than with low value capacitors. The reason for this is that, larger capacitors have larger plate areas that are closer together. Therefore, their dielectrics must have large areas and be thin. Recall, resistance reduces as the cross-sectional area is increased or when the length or thickness is decreased. So, larger the capacitor, lower the leakage resistance, and higher is the leakage current.

Normal leakage resistance across a good capacitor has to be very high. Depending upon the type of dielectric used, the normal resistance varies. For paper, plastic, mica and ceramic capacitors the normal resistance will be of the order of 500 to 1000 M Ω or more. For electrolytic capacitors the normal resistance will be of the order of 200 K Ω to 500 K Ω .

A capacitor is said to have become leaky when the resistance across it is less than normal when read with any average quality ohmmeter.

Checking capacitors

The two simple methods to check a capacitor is by carrying out,

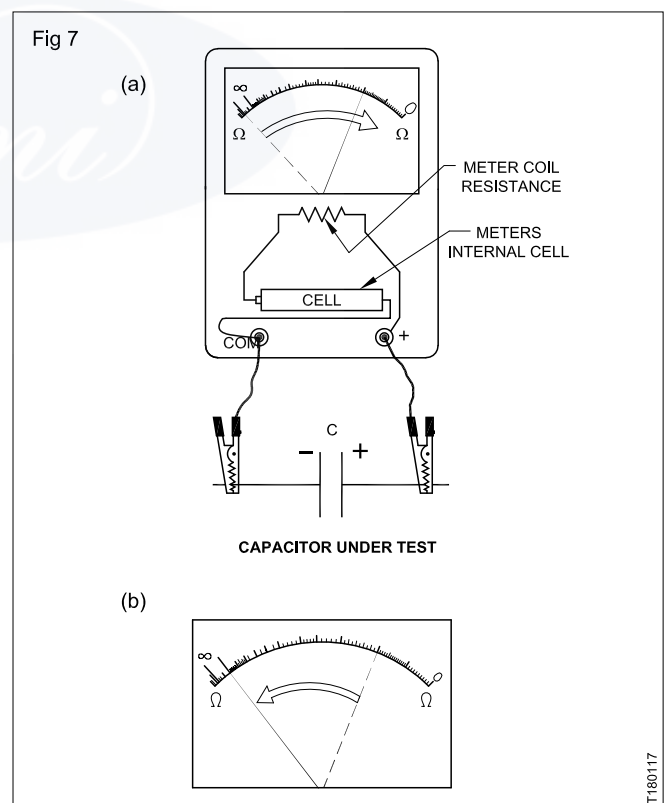
- capacitor action-normal resistance test, using a ohmmeter/multi-meter (This test is also referred as quick-test)
- charging-holding test, using a battery and voltmeter/multi-meter.

Capacitor action-normal resistance test

When an ohmmeter is connected across a fully discharged capacitor, initially, the battery inside the meter charges the capacitor. During this charging, at the first instance, a reasonably high charging current flows. Since more current through the ohmmeter means less resistance, the meter pointer moves quickly towards zero ohms of the meter scale as shown in Fig 7a.

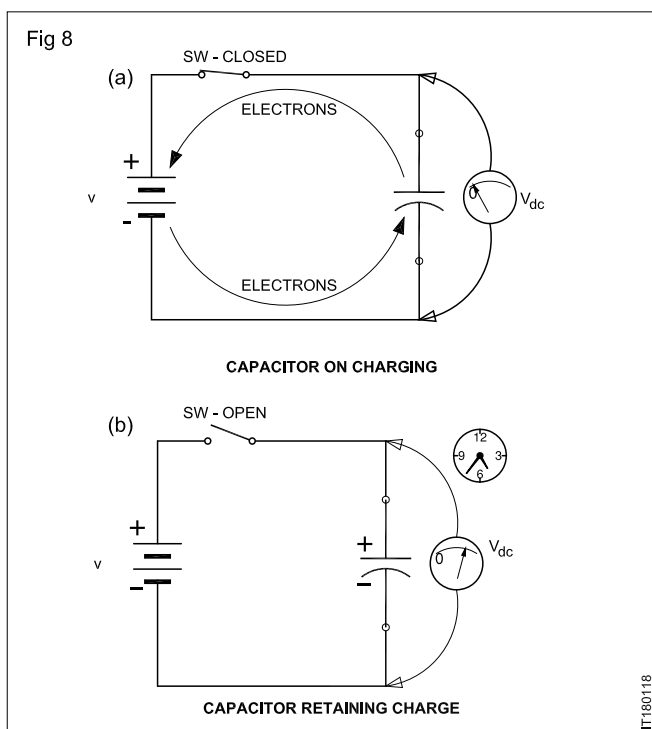
After the initial charging, the charging current to the capacitor slowly decreases (as the voltage across the capacitor increases towards the applied voltage). Since less and less current through the ohmmeter means high and higher resistance, the meter pointer slowly moves towards infinite resistance of the meter scale as shown in Fig 7b. Finally, when the capacitor is completely charged to the ohmmeter internal battery voltage, the charging current is almost zero and the ohmmeter reads the normal resistance of the capacitor which is a result of just the small leakage current through the dielectric. This changing effect, commonly known as *Capacitor action* indicates, whether the capacitor can store charge, whether the capacitor is excessively leaky, whether the capacitor is fully short-circuited or whether the capacitor is fully open-circuited.

The capacitor-action test is most suitable for high value capacitors and specially electrolyte capacitors. When small value capacitors such as ceramic disc or paper capacitors are tested for capacitor-action, due to the extremely low charging current the capacitor-action cannot be observed on the meter dial. For such small value capacitors the capacitor-charging-holding test is preferred. However if small capacitors are subjected for capacitor-action test, if the meter shows high resistance the capacitor can be taken as not shorted and hence may be taken as good.



Charging-holding test on capacitors

In this test, a given capacitor is charged to some voltage level using an external battery as shown in Fig 8a.



Once the capacitor is charged to the applied voltage level, the battery is disconnected and the voltage across the capacitor is monitored as shown in Fig 8b. The voltage is monitored for a period of time to confirm whether the capacitor is able to hold the charge atleast for a small period of time (of the order of few seconds).

In this test, when the capacitor is tried for charging, if the capacitor does not charge at all even after connecting the battery for a considerable period of time, it can be concluded that the capacitor is either short-circuited or fully open circuited.

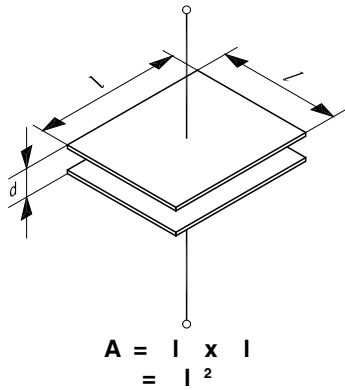
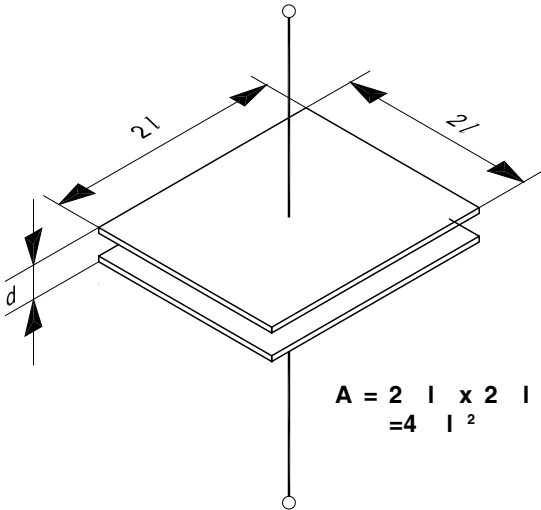
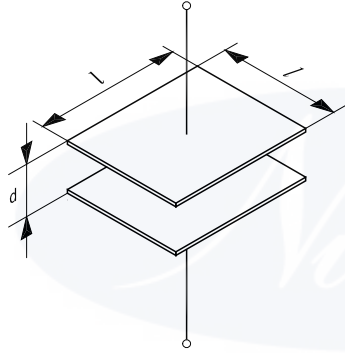
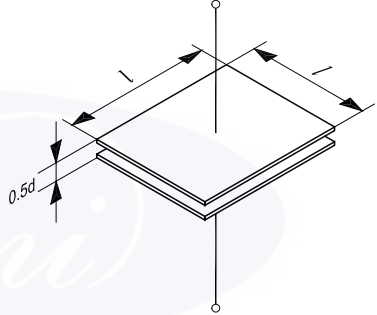
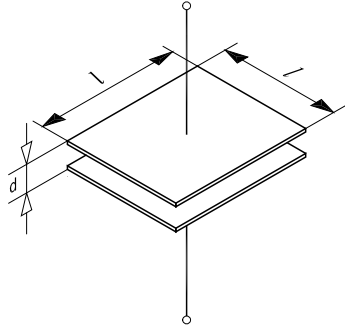
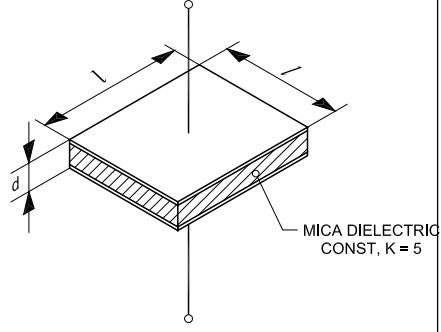
If the capacitor is unable to hold the charge even for a considerably small period of time, then it can be concluded that the capacitor is excessively leaky.

The following points are important and are to be noted to get correct results from this test :

- 1 If the capacitor to be tested is marked with + and - at its terminals (polarised-capacitor) then connect the battery with the same polarity. If a polarised capacitor is tried for charging with opposite polarity, the capacitor may get permanently damaged.
- 2 Use a FET input voltmeter or high ohm/volt voltmeter to monitor the holding of voltage across the charged capacitor. This is because a low ohm/volt voltmeter will draw current from the charged capacitor resulting in the early discharge of stored charges on capacitor.

The term FET stands for a type of transistor discussed in subsequent units. A FET input voltmeter is a high quality voltmeter having very high ohms/volts. This meter draws almost zero current while measuring voltage across any two terminals. Other average voltmeters draw current in the range of a few hundreds of micro-amps to a few milli-amps while measuring voltage.

CHART - 1
Physical dimensions versus Capacitance value

FACTOR	LESS CAPACITANCE	MORE CAPACITANCE	FACTOR
AREA OF CROSS SECTION (A) OF PLATES	 $A = l \times l$ $= l^2$	 $A = 2l \times 2l$ $= 4l^2$	DOUBLING THE PLATE AREA DOUBLES THE CAPACITANCE
DISTANCE BETWEEN PLATES			REDUCING THE DISTANCE BETWEEN PLATES BY HALF DOUBLES THE CAPACITANCE.
TYPE OF DIELECTRIC MATERIAL		 MICA DIELECTRIC CONST, K = 5	CAPACITANCE INCREASES AS MANY TIMES AS THAT OF INCREASE IN VALUE OF k

Main types of fixed value capacitors

Type	Main sub-types	Dielectric used	Construction	Available capacitances	Rated voltage	Applications
Paper	Foil type &	Impregnated special craft paper	Rolled foils	0.001-1mF	200-1600VDC	Motor - starting, PF correction power supply - filters.
	Metalized type	Special tissue paper				
Plastic film	Foil type &	i) Polystyrene ii) Polyester (Mylar capacitor) -	- Rolled foils	0.001-1mF 0.001-1mF 0.001-0.47mF 0.01-18mF 4-60mF 0.01-10mF	100-200VDC- 100-600VDC- 400-1600 VDC- 50-200VDC- 400 VAC, 50 Hz 100-600 VDC	General purpose, high stability. General purpose. RF circuits. General purpose. AC motors. Coupling, RF filtering.
	Metalised type	iii) Poly propylene- iv) Poly carbonate v) Metallized polypropylene- vi) Metallized polyester- vii) Polystyrol (Styroflex)				
Ceramic	Disc type	Class-1 (Non ferro-electric) -Steatite (Talc) -Mix of MgO, TiO ₂ -TiO ₂ , CaO	Drawn ceramic films Moulded tubes Substrate- Screening-sintering	1PF -1mF 1PF-1000PF 1pF-10mF	50-500VDC 500-5KVDC 50-200VDC	General purpose, RF. General, VHF. VHF, RF coupling. Coupling in VHF range. Decoupling in VHF range. HF circuit feeders.
	Tube type Monolithic (chip type)	Class-2 (Ferro-electric) -Barium titanate -Ba, Sr, TiO ₂ +Mg, Zr				
	Feed-through- stand-off- button type					
Electro-lytic	Aluminium (polar, non-polar)(Wet, dry type)	Aluminium oxide	Rolled foil - metallic can Rolled foil - Can/cup/tank	1-500,1000mF 0.1-1000mF	5-500 VDC 3-125VDC	Power supplies, filters. Space electronics. Nonpolar Al and Ta capacitors are used in loudspeaker cross-overs.
	Tantalum (polar, non-polar)(Wet, dry type)	Tantalum pentoxide				
Mica	Stacked mica- Silvered mica Button type	White mica, Rose mica, Amber mica	Stacked	5 pF-10,000pF 5pF-3300pF	50-100 50-500	High frequency High frequency H.F line feeders
		Thin layer of glass	Stacked	5 pF-5000pF		VHF applications
Vitreous Enamel		Mixture of silica, Potassium, lead oxide and fluorides	Deposited in layers		50-500	

CHART - 3 : Physical appearance of types of fixed value capacitors

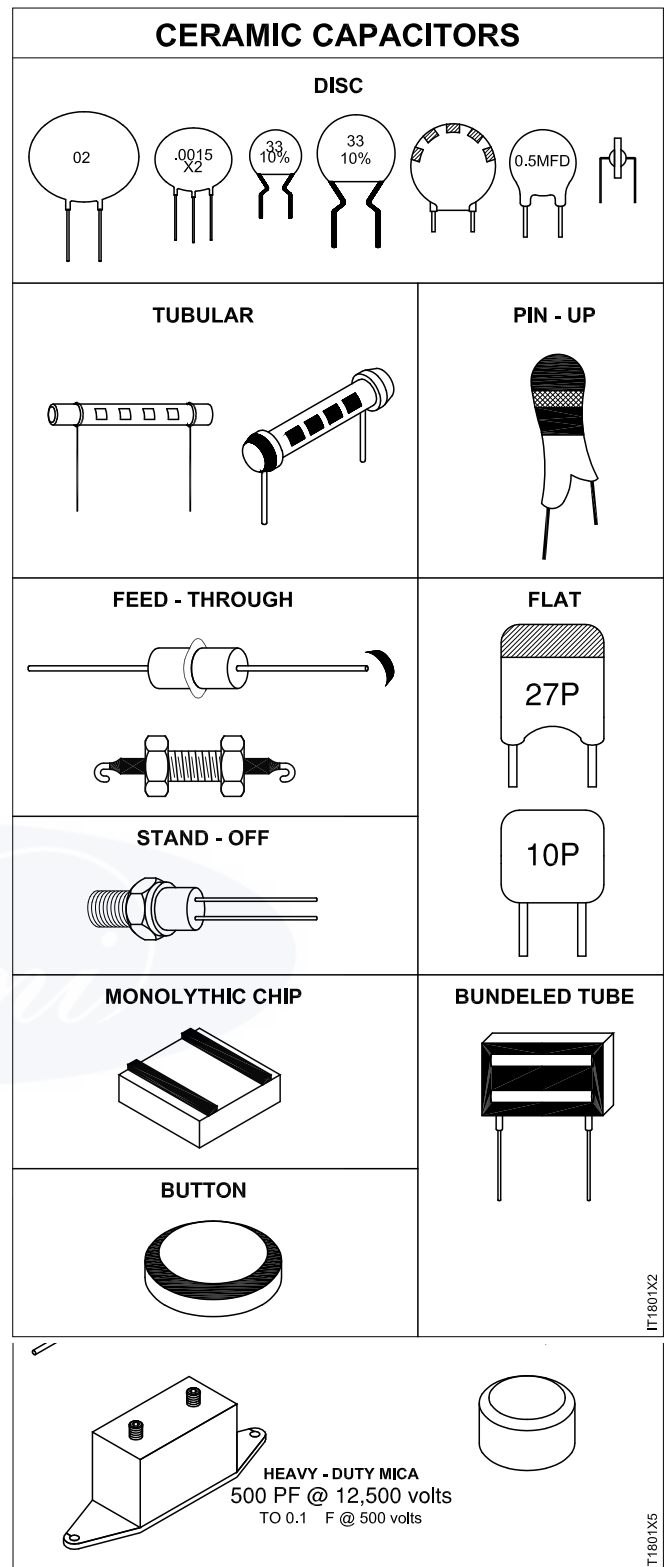
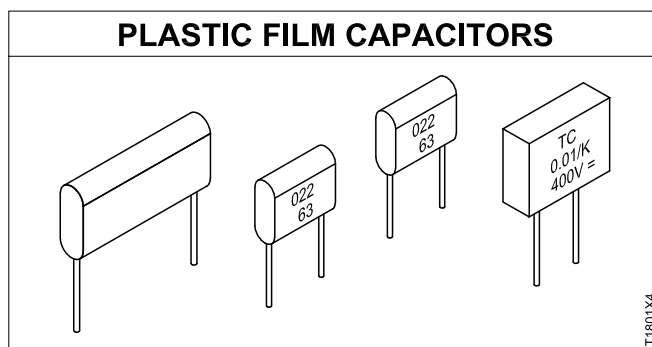
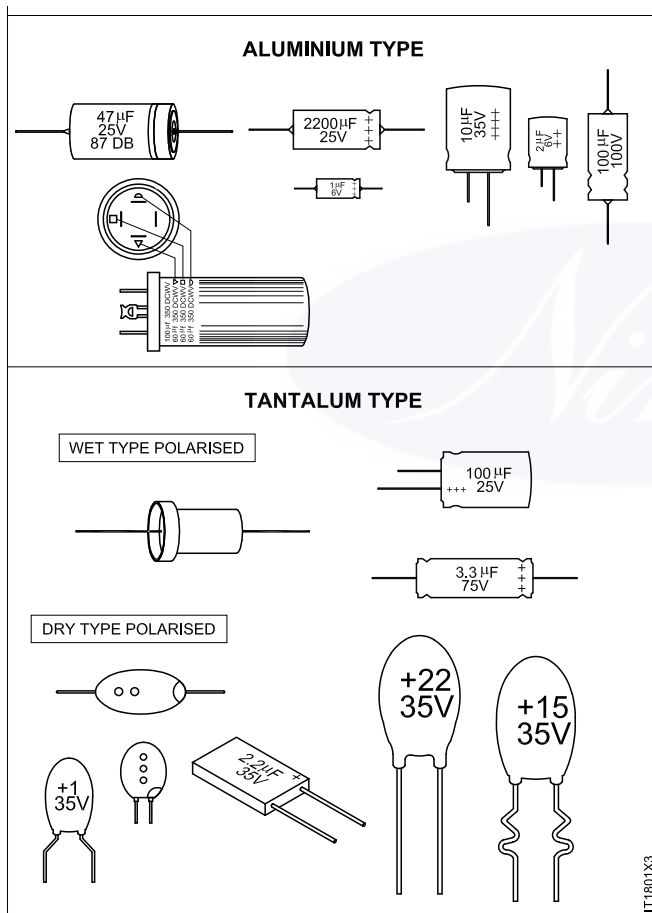
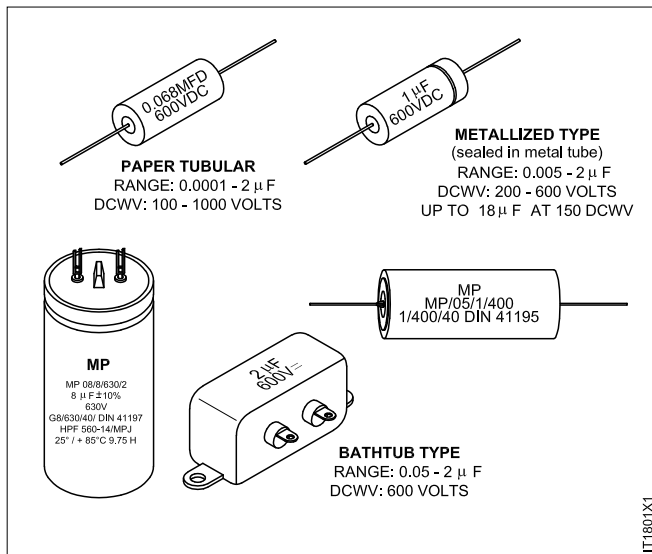
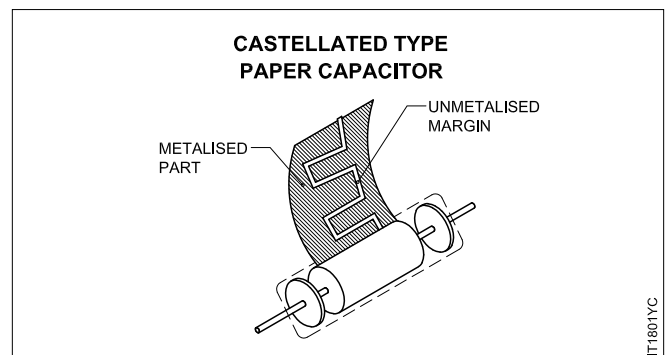
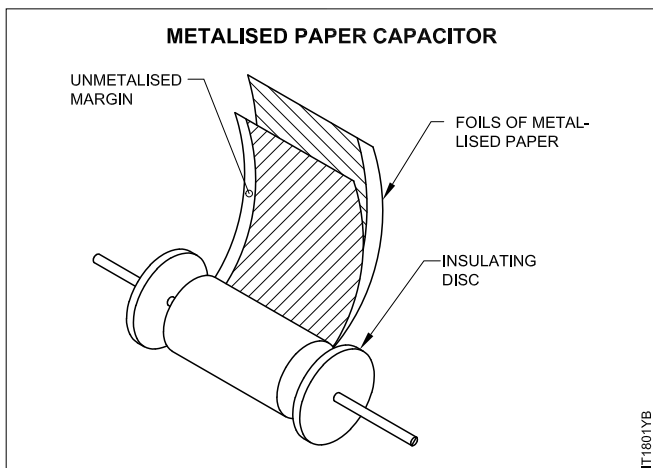
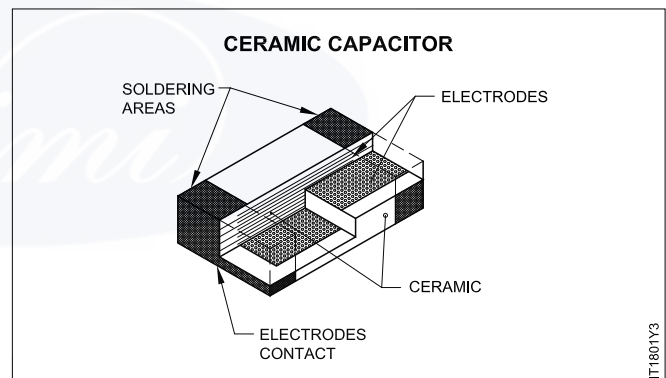
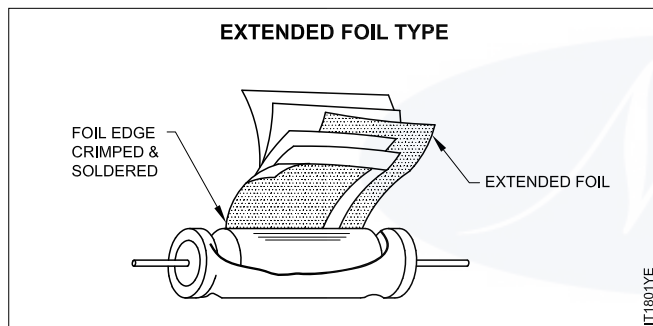
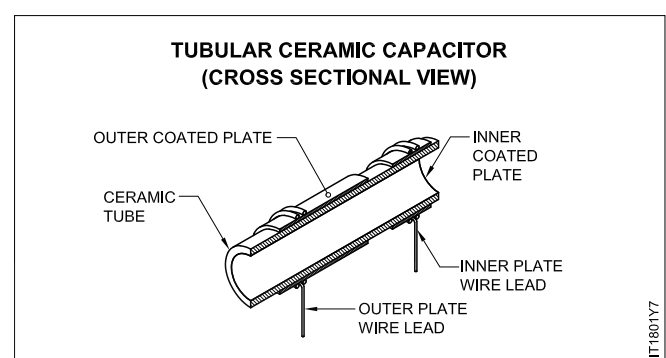
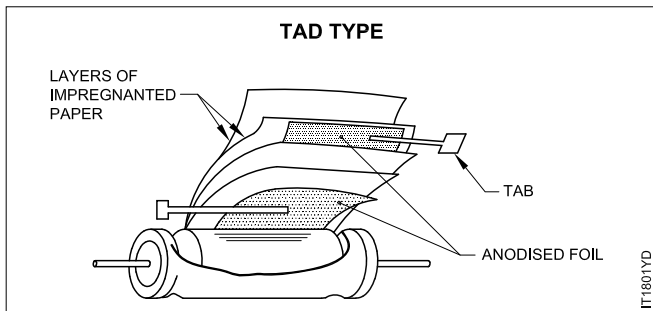
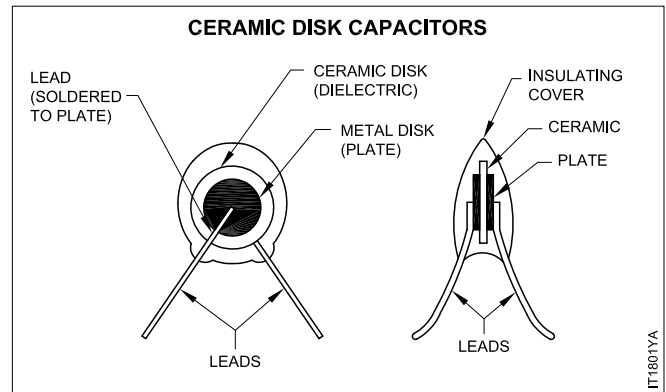
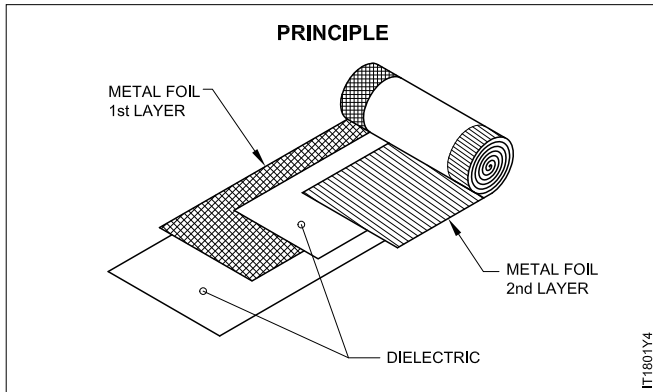
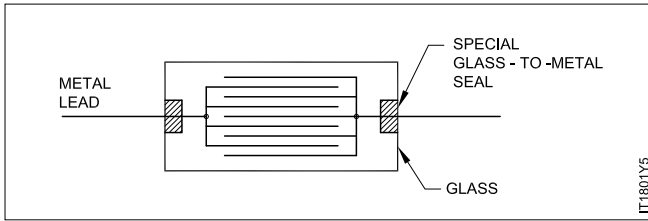


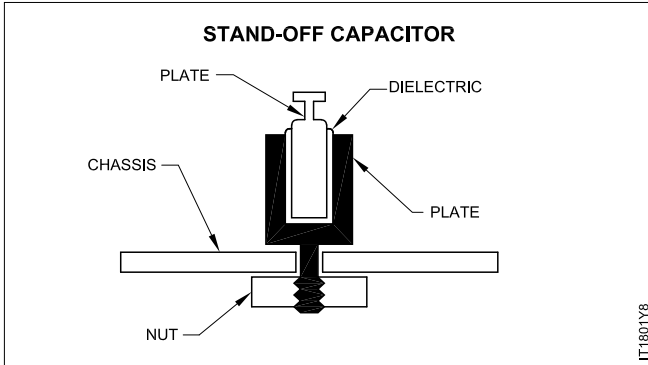
CHART - 4 CONSTRUCTIONAL DETAILS OF FIXED VALUE CAPACITORS



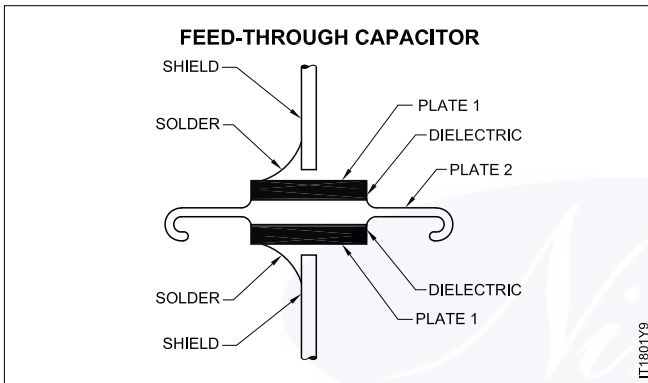
GLASS CAPACITORS



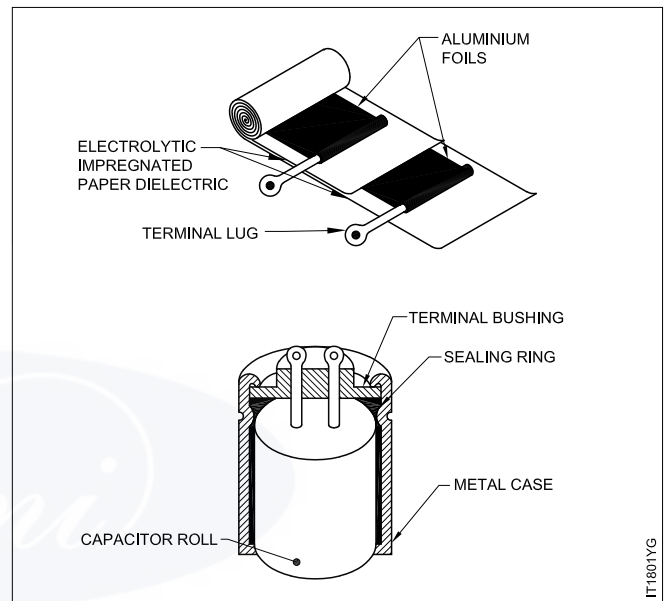
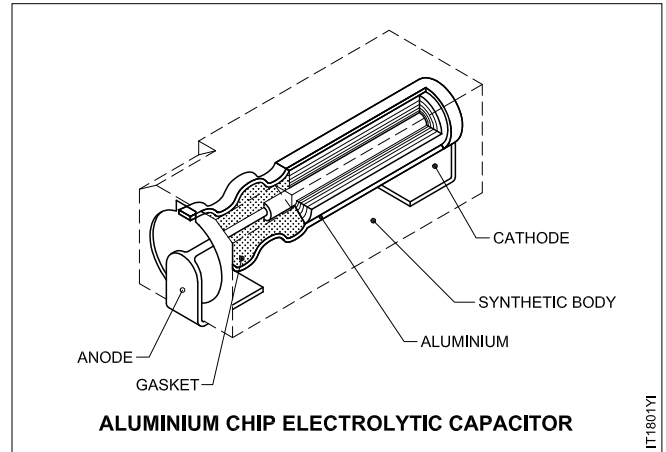
STAND-OFF CAPACITOR



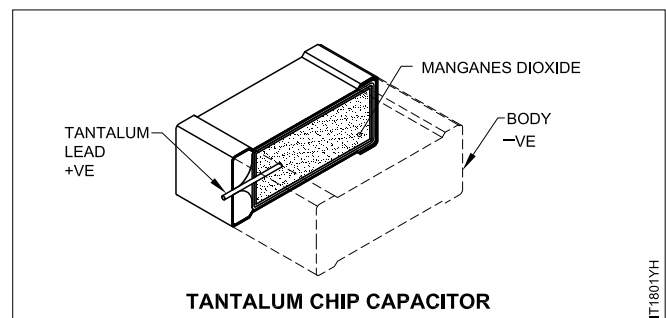
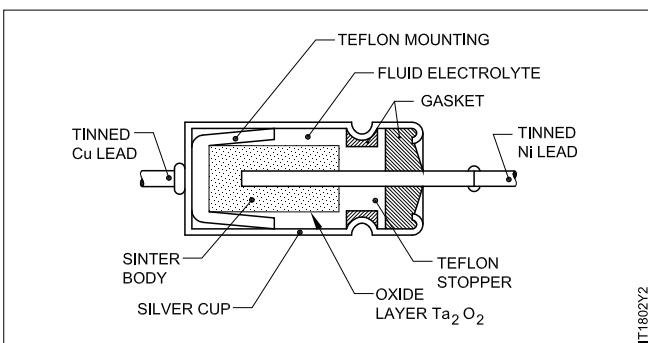
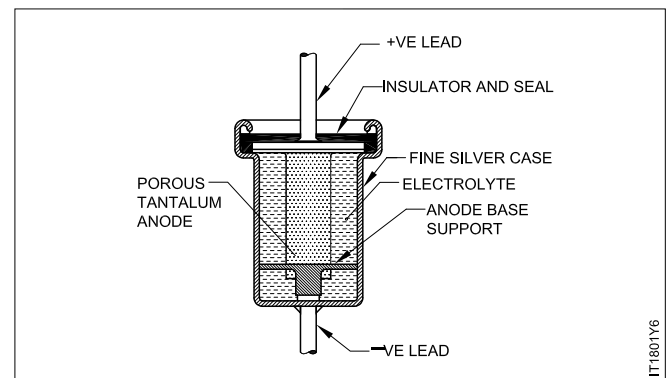
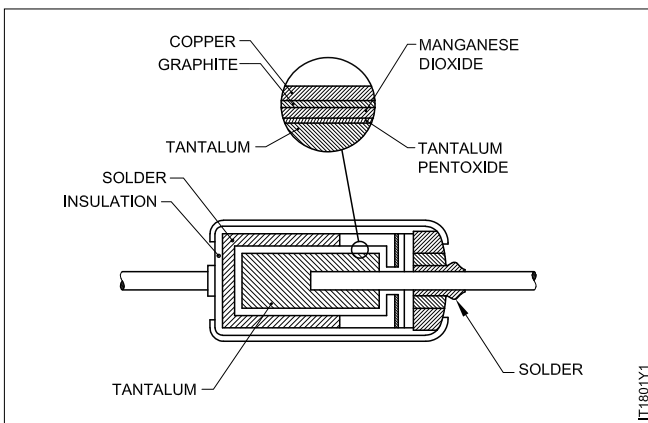
FEED-THROUGH CAPACITOR



ELECTROLYTIC CAPACITORS ALUMINIUM TYPE



TANTALUM CAPACITORS



TANTALUM CHIP CAPACITOR

Measurement of capacitive Reactance at different frequencies

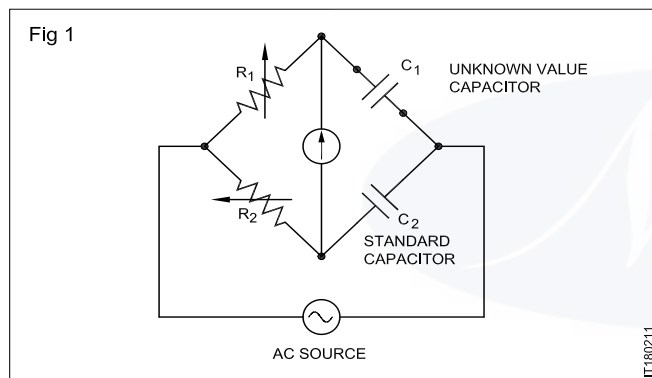
Objectives: At the end of this lesson you shall be able to

- state methods of measuring capacitance value
- explain how capacitors block dc signal and allow ac signals
- list a few applications of capacitor
- state the relationship between current and voltage in a pure capacitive circuit
- state the effect of frequency and capacitance on capacitor current
- explain the term capacitive reactance and effect of R in series with a capacitor.

The capacitance value of capacitors can be measured using an AC bridge or digital meters specially designed to measure capacitance values.

AC bridge method of measuring capacitance

There are several methods of measuring capacitance using AC bridges. The easiest of all these is the *De Santy method*. This method adopts the technique of comparing the value of an unknown capacitor with that of a known capacitor as shown in Fig 1.



C_1 = Capacitor whose capacitance is to be measured.

C_2 = Known standard capacitor

R_1 and R_2 = Non-inductive resistors

The bridge is balanced by varying either R_1 or R_2 or both. With the bridge in a balanced condition,

$$\frac{R_1}{R_2} = \frac{C_2}{C_1}$$

Therefore the value of the unknown capacitor C_1 can be calculated from the known values of R_1 , R_2 and C_2 as follows:

$$C_1 = C_2 \frac{R_1}{R_2}$$

NOTE: A perfect balance of bridge is not possible if the capacitor under test is not free from dielectric loss.

For details of other bridge methods used for capacitance measurement, refer reference books listed at the end of this book.

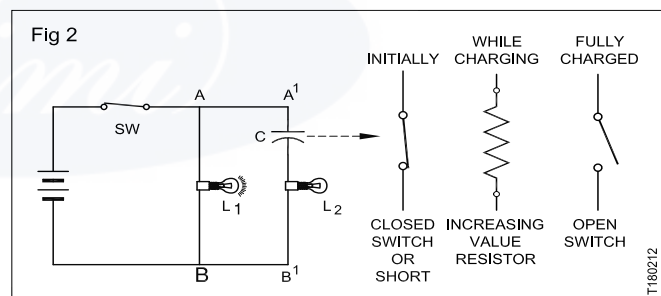
Digital meters for capacitance measurement

Digital meters for exclusive measurement of capacitance are available commercially. These meters can measure values from a few hundred pico-farads to a few thousands of micro-farads.

Digital meters are also available which can measure resistance, inductance and capacitance, known as RLC meters. These meters, in addition to measuring the capacitance values, can also measure other properties of capacitors such as leakage current, insulation resistance.

Capacitors in DC circuits

Consider a test circuit shown in Fig 2.

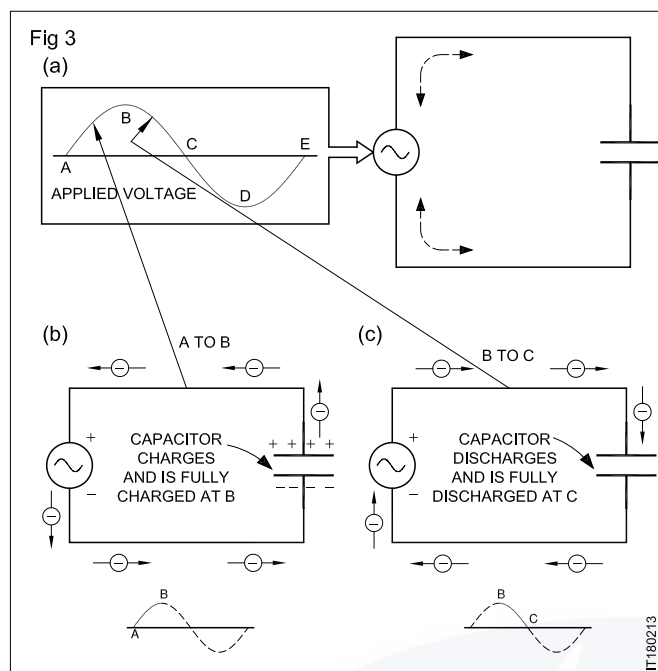


When switch SW is put ON, lamp L_1 glows continuously whereas lamp L_2 glows brightly only momentarily, and then the brightness becomes dim and eventually L_2 glows no more. The reason is, capacitor C in series with L_2 , initially behaves as a closed switch. During this time a relatively large current (limited by the resistance of the lamp) flows through the path $A^1 - B^1$ of the circuit. This current starts charging the capacitor. As the capacitor slowly gets charged towards the voltage of battery supply, the voltage built-up across the capacitor opposes the battery voltage. This then makes the capacitor appear as an increasing value resistor. Once the capacitor gets fully charged to the level of the battery voltage, the charged capacitor voltage opposes the battery voltage, and hence, current through the path $A^1 - B^1$ becomes zero. This makes the capacitor appear as an opened switch.

From this behavior of capacitor in a DC circuit, it can be said that *capacitor blocks DC*.

Capacitor in AC circuits

As in a DC circuit, a capacitor also gets charged by an AC voltage. But, as the polarity of AC voltage reverses periodically, the charge on the capacitor cannot remain stable as in a DC circuit. Consider a circuit as shown in Fig 3 where a sine wave voltage from an AC source is connected across a capacitor.



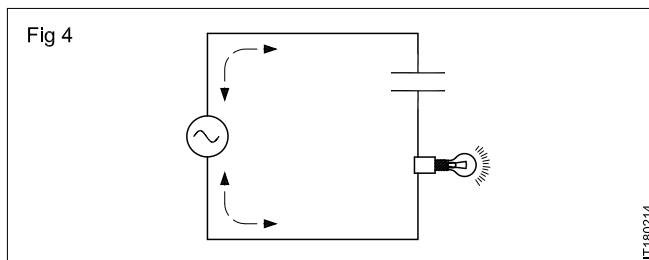
The AC charging voltage starts at zero amplitude (a in Fig 3) and increases in a positive direction. This increase in voltage charges the capacitor. When the charging AC voltage reaches its peak value (b in Fig 3), maximum voltage is built-up in the capacitor, and, hence, the charging current becomes zero.

Having reached its peak value at the end of the first quarter cycle (90°) shown in Fig 3b, the charging AC voltage begins to decrease in amplitude. At this instant, the voltage built up across the capacitor exceeds the value of the charging voltage source. Since the source and capacitor are in parallel, to maintain equal voltage, the capacitor starts discharging current into the voltage source as shown in Fig 3c. This process of discharging continues and when the charging source reaches point C, the voltage across the capacitor also would have reduced to zero.

The charging, discharging of capacitor voltage will repeat during the negative half cycle. However, the polarity change in the charging voltage produces a change in the polarity of the voltage across the capacitor.

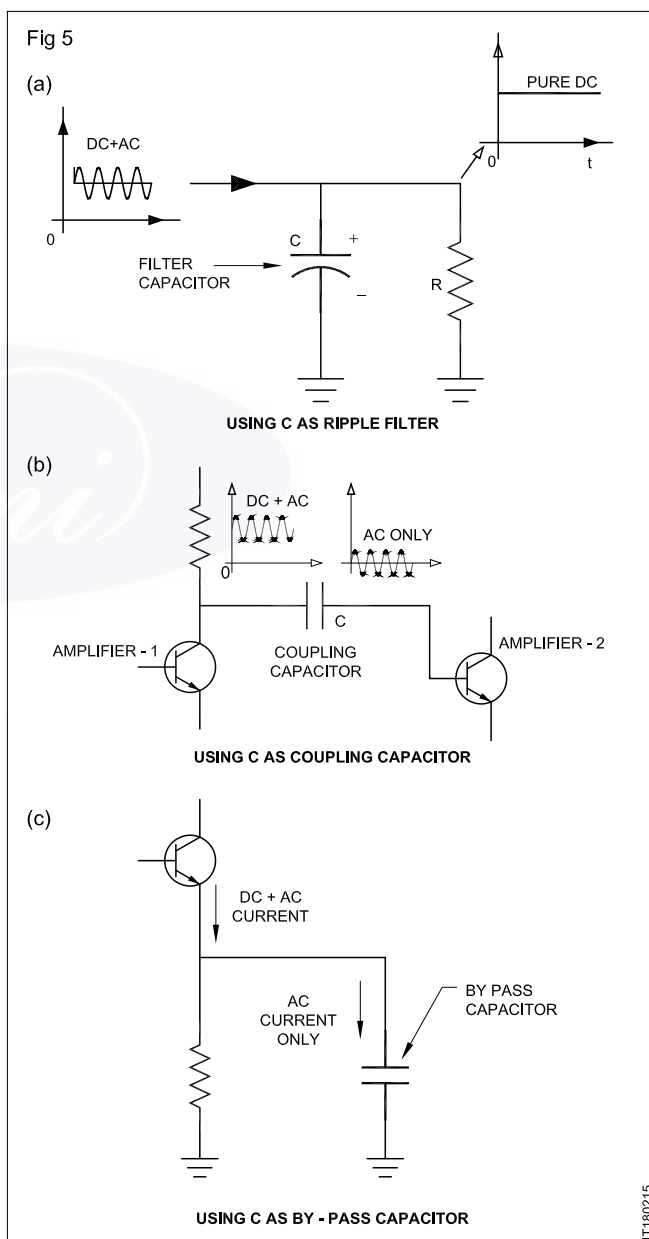
From the above it is evident that the variations in charge on the plates of the capacitor also follows the form of the sine wave.

Fig 4 shows a lamp connected in series with a capacitor. In this circuit the lamp continues to glow with normal brightness as long as the supply exist.



From this behavior of capacitor in an AC circuit, we can say that the *capacitor allows AC to pass*.

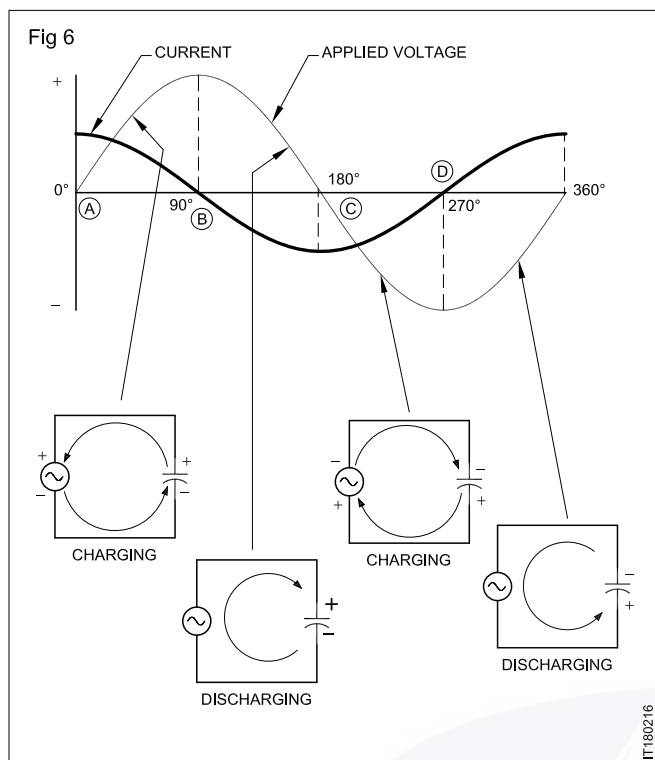
This property of capacitors to block DC voltage but pass AC voltage is made use of in the application of capacitors for coupling, bypassing and filtering of AC signals as illustrated in Fig 5a, 5b and 5c.



[H.I: Instructor should discuss the circuits given in Fig 5 and discuss the function of capacitor in each circuit at Fig 5.]

Voltage and current relationships in capacitors

As shown in Fig 6, when an AC voltage source is connected across a capacitor, maximum current flows in the circuit the instant the source voltage begins its sinusoidal rise from zero.



At first it may seem strange that the maximum current flows when the source voltage is at its lowest value. But remember, this current is actually the movement of free electrons from the negative terminal of the source to one capacitor plate, and from the other plate to the positive terminal of the source. At the first instant of charging, since the plates are neutral, they present no opposition to the source voltage. Therefore, by Ohm's Law, if the opposition to current flow is very low, even a small applied voltage can cause considerable current to flow in the circuit.

As the source voltage rises, the charges on the capacitor plates build up due to the current flow. This built up voltage across the capacitor, then, presents an increasing opposition to the source voltage and so the current decreases. When the source voltage reaches its peak value, the charged voltage across the capacitor plates is also maximum. This charge is sufficient to cancel the source voltage completely, and so the current flow in the circuit becomes zero as shown in Fig 6 (point B in the graph). As the source voltage begins to decrease, the charge on the capacitor plates becomes greater than the voltage of the source. So the capacitor starts discharging.

Thus, at point B of Fig 6 the applied voltage passes through its maximum value and begins decreasing while the current in the circuit passes through zero and changes direction. As can be seen from the graph (Fig 6), this results in a 90-degree phase difference between the source voltage and current, with the current leading the applied voltage by 90°. This 90-degree difference is maintained throughout the cycle of the applied voltage.

Due to this 90° leading current, when the applied voltage is passing zero, the circuit current will be at its maximum and when the applied voltage is maximum, the circuit current will be zero.

Therefore, in a capacitor, *the current through the capacitor leads the applied voltage by 90 degrees or the applied voltage lags the circuit current by 90 degrees.*

Effect of frequency on capacitor current

The amount of current flowing in any circuit is nothing but the number of electrons that flow past any point in the circuit, in a unit time. In other words, current is equal to the rate of flow of electrons. As an equation, this can be written as,

$$I = \frac{Q}{t} \text{ amperes}$$

where,

Q = number of electrons being transported in time t .

Current flow in an AC capacitive circuit is made up of free electrons flowing between the terminals of the voltage source and the plates of the capacitor. The electrons flow in one direction while the capacitor is charging, and in the reverse direction when the capacitor is discharging. As shown in Figs 7a and 7b, for any given capacitor, how fast a capacitor acquires maximum charge depends on the frequency of the applied voltage.

This is because the charge on a capacitor follows the applied voltage exactly. Higher the frequency of AC, the faster the applied voltage reaches its maximum voltage. So, it follows that higher the frequency, the faster the capacitor will reach its peak charge. So, when two identical capacitors are connected to identical voltage sources, which have different frequencies, as shown in Fig 7, although both capacitors will finally have the same value of Q , capacitor C_1 will have it earlier than capacitor C_2 .

Due to this time difference in reaching the same final value of Q , at any instant of time currents through the two capacitors will be different as shown in the graph at Fig 7c. So, higher frequency causes greater current, and the lower frequency causes lesser current to flow through the capacitor.

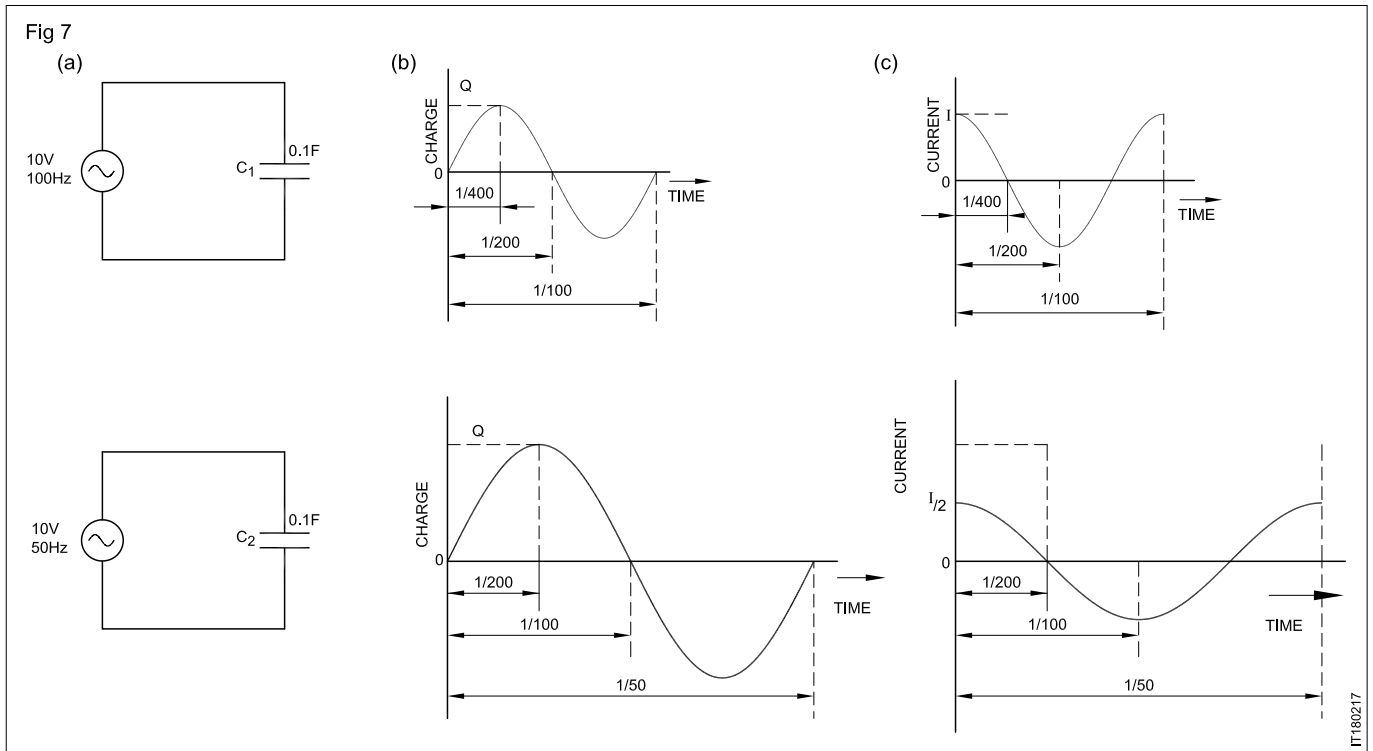
Summarizing, *current through a capacitor is directly proportional to the frequency of the applied voltage.* High frequencies result in large currents, whereas low frequencies produce small currents in the same capacitor.

Example: For the circuits shown in Fig 7, calculate the maximum current that will flow in each circuit if both capacitors would reach a maximum charge of 1 coulomb.

SOLUTION

a) For the circuit with 50 Hz, it would take 1/200 of a second to reach a charge of 1 coulomb.

Therefore, using the formula $I = Q/t$



$$I = \frac{1}{1/200} = 200 \text{ amperes.}$$

b) For the circuit with 100 Hz, it would take 1/400 of a second to reach a charge of 1 coulomb.

Therefore, using the formula $I = Q/t$

$$I = \frac{1}{1/400} = 400 \text{ amperes.}$$

Effect of capacitance on capacitor current

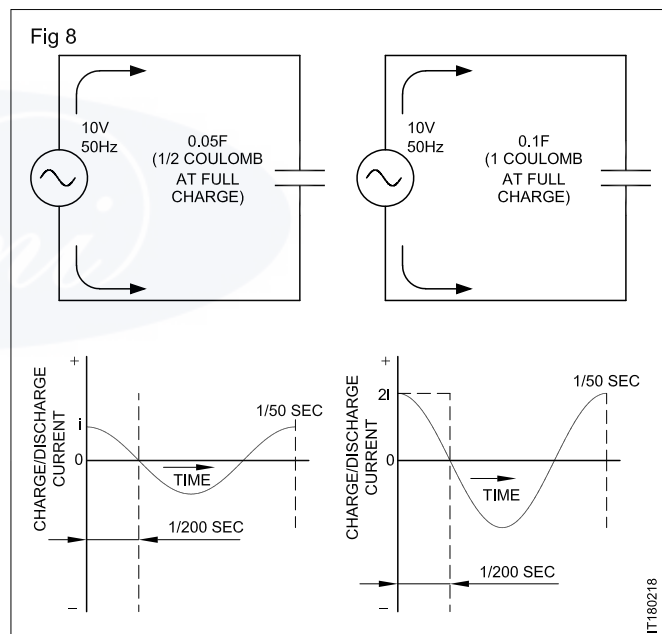
Recall that capacitance is the amount of charge stored in a capacitor for each volt applied across the capacitor ($C = Q/V$). So, the total charge stored is equal to the capacitance times the applied voltage or $Q = CV$.

Assume two circuits having different values of capacitors. The applied voltages to both the circuits have the same amplitude (V) and same frequency as shown in Fig 8.

The charges stored in the capacitors in each circuit will be different because $Q = CV$. The capacitor with a larger capacitance will store a larger charge because of the larger value of C. This means that with larger capacitance, more free electrons will flow during the charge and discharge cycles. If the time required for fully charging the two capacitors were the same, which would be the case if the applied voltages had the same frequency, the current through the capacitor with the larger capacitance would be greater than the current through the smaller capacitance capacitor.

From this it can be concluded that, *the current through a capacitor is directly proportional to the capacitance.*

In the circuits in Fig 8, although the amplitude and frequency of the applied AC voltage is the same to both the circuits,



In the first circuit with $C = 0.05\mu F$, $Q = C.V = 0.5$ coulomb.

The capacitor will build up a charge of 1/2 coulomb in 1/4 of a cycle which is 1/200 sec.

Using the equation $I = Q/t$,

the maximum current will be 100 amperes.

In the second circuit with $C = 0.1\mu F$, $Q = C.V = 1$ coulomb.

The capacitor will build up a charge of 1 coulomb in 1/4 of a cycle which is again 1/200 sec.

Using the equation $I = Q/t$,

the maximum current will be 200 amperes.

From the above example it is clear that, by doubling the value of C the charging/discharging current through the circuit also doubles.

Capacitive reactance

Similar to resistors and inductors, a capacitor also offers opposition to the flow of AC current. This opposition offered to the flow of current by a capacitor is called **capacitive reactance** abbreviated as X_C .

Recall expressions,

$$I = \frac{Q}{t} \text{ and } Q = CV$$

Substituting $Q = CV$ in $I = Q/t$

$$I = \frac{CV}{t}$$

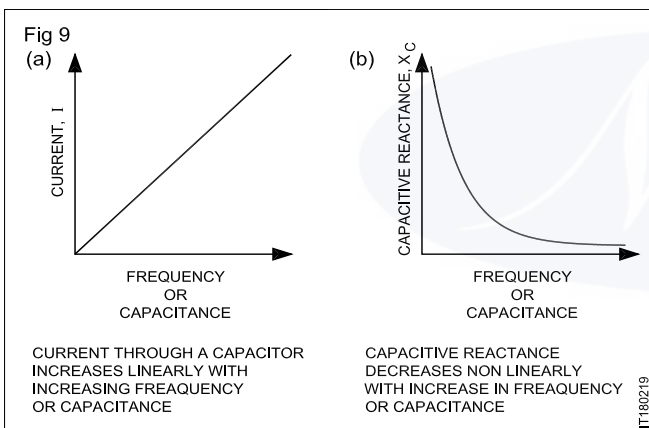
This means,

$I \propto C$, $I \propto V$ and $I \propto f$ (Because, $1/t = f$)

From the above equation, the amount of AC current that a capacitor conducts depends on;

- the frequency (f) of the applied voltage
- the capacitance (C) of the capacitor
- the amplitude of the applied voltage (V).

Fig 9a shows the graph of variation of current (I) through a capacitor with frequency or capacitance when the applied voltage is kept constant.



Since current flow through a capacitor is directly proportional to frequency and capacitance, the opposition to current flow by the capacitor is inversely proportional to these quantities.

Capacitive reactance, X_C can be mathematically represented as;

$$X_C = \frac{1}{2\pi fC} \text{ ohms}$$

where

X_C is the capacitive reactance in ohms

f is the frequency of the applied voltage in Hz

and C is the capacitance in farads.

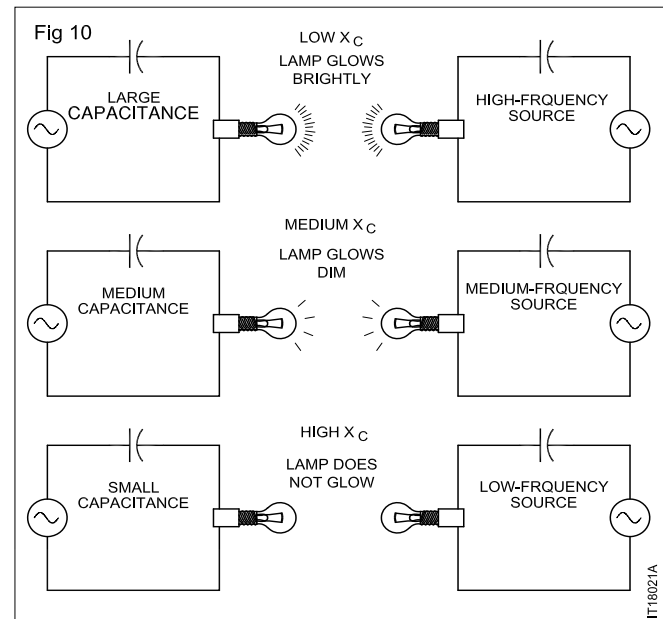
Fig 9b shows the graph of variation of X_C with frequency or capacitance.

Capacitive reactance, X_C , expressed in ohms, acts just like a resistance in limiting the AC current flow. When the capacitive reactance is known, current can be calculated

using the formula;

$$I = \frac{V}{X_C} \text{ amperes}$$

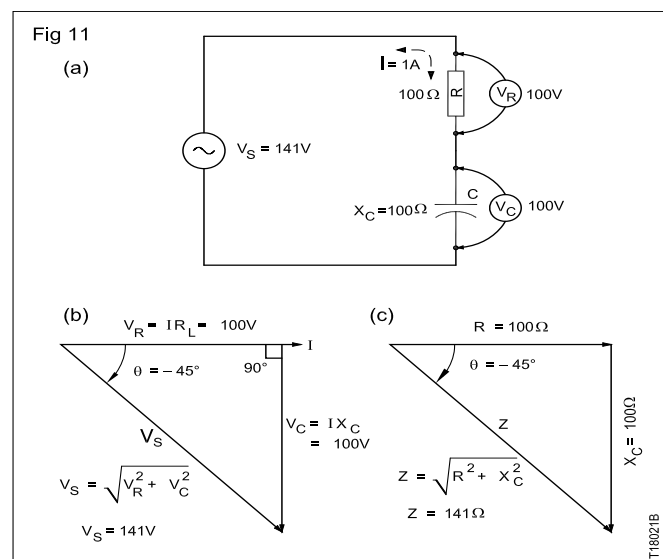
Fig 10 illustrates the effect of capacitive reactance X_C in a simple AC lamp circuit.



Series resistor - Capacitor circuit

When resistance is connected in series with a capacitor as shown in Fig 11, the current in the circuit is determined both by the resistor R and the capacitive reactance X_C of the circuit. Current I is the same in X_C and R since they are in series. Each component has its own series voltage drop, equal to I_R for the resistance and IX_C for the reactance.

If only the resistance R in the circuit at Fig 11 is considered, the I_R voltage has the same phase as I , because resistance provides no phase shift. If the capacitive reactance X_C alone is considered, since the current through a capacitor leads the voltage by 90° , its voltage drop lags the series current I by 90° .



In the phasor diagram at Fig 11b, the resistive voltage drop I_R is shown in phase with I . The capacitor voltage IX_C is shown at -90° , (clockwise from I and IR) as the capacitive voltage lags behind current I by 90° .

In phasor representation, remember that positive angles are measured counter-clockwise and Negative angles are measured in clockwise direction.

From the phasor diagram at Fig 11b, it can be seen that voltages V_R and V_C are out of phase by 90° . Therefore, the total voltage V_s across both R & X_C is given by,

$$V_s = \sqrt{V_R^2 + V_C^2}$$

where,

V_R = Voltage across resistor R in volts
 V_C = Voltage across capacitor C in volts.

Impedance Z of a series R-C circuit

The voltage triangle in Fig 11b corresponds to the impedance triangle in Fig 11c. This is, because the common factor I with both X_C and R gets cancelled. The phasor sum of X_C and R is called the *impedance* Z of the circuit.

From the Fig 11b, Z is given by;

$$Z = \sqrt{R^2 + X_C^2} \text{ ohms}$$

With R and X_C in ohms, Z is also in ohms.

In Fig 11c, the values of Z is equal to

$$Z = \sqrt{100^2 + 100^2} \text{ ohms}$$

$$Z = 141 \text{ W.}$$

This total impedance of 141W, produces a current of 1 A in the series circuit ($I=V/Z$).

The voltage drop across R is $I_R = 1 \text{ A} \times 100 \text{ ohms} = 100\text{V}$.

The voltage drop across C is $IX_C = 1 \text{ A} \times 100 \text{ ohms} = 100\text{V}$.

The arithmetic sum of voltage drops across R and X_C will be more than the applied voltage if the phase angle is not considered.

The phasor sum of the two series voltage drops of 100V each equals the applied voltage of 141V as shown in Fig 11b. Also, the applied voltage is equal to $I \times Z$, or 1×141 , which is 141V.

Phase angle q in series RC circuit

As shown in Fig 11b, the phase angle q , is the angle between the source voltage V_s and the series current I in a series R-C circuit. As shown in Fig 11b and 11c, the value of q can be found either by plotting a graph of V_R , V_C and measuring the angle or can be calculated from the impedance triangle.

In a series R-C circuit, the phase angle q is negative. This is because, voltage across X_C lags the current. To indicate the negative phase angle, the 90° phasor line is pointed downward from the horizontal reference as shown in Fig 11b.

Phase angle q can also be calculated knowing the values of R and X_C as follows;

From the voltage or impedance triangle, shown in Fig 11c,

$$\tan \theta = - \frac{X_C}{R} \text{ or } \theta = \tan^{-1} \left(- \frac{X_C}{R} \right)$$

Using this formula for the circuit in Fig 11a,

$$\tan \theta = - \frac{X_C}{R} = - \frac{100 \Omega}{100 \Omega}$$

$$\tan q = -1$$

$$\text{Therefore, } q = \tan^{-1} (-1) = -45^\circ.$$

The negative sign attached with q indicates that the applied voltage V_s lags behind the circuit current I , by angle q .

Example: A Resistance of 40W is in series with a capacitor having a capacitive reactance of 40W. Calculate the Z and q of the circuit.

SOLUTION:

$$Z = \sqrt{R^2 + X_C^2} = \sqrt{(40)^2 + (40)^2}$$

$$Z = \sqrt{1600 + 1600} = \sqrt{3200}$$

$$Z = 56.57 \text{ W.}$$

$$\tan \theta = - \frac{X_C}{R} = - \frac{40 \Omega}{40 \Omega} = -1$$

$$\text{or } q = \tan^{-1} (-1) = -45^\circ.$$

From the above example, it can be seen that, when the value of R and X_C are equal, q is midway of -90° . The higher the X_C compared with R , the circuit becomes more and more capacitive and q increases towards -90° . In such highly capacitive circuits, the current leads the applied voltage heavily. If R in such a circuit is reduced to zero, the entire applied voltage is across X_C , and q equals -90° .

Note that if the $C:R$ ratio in a circuit is 10:1 or more then the circuit will be practically all capacitive. The phase angle will be almost -90° and the total impedance Z will be approximately equal to X_C itself.

On the other hand, when R is ten times more than X_C , the series circuit will be mainly resistive. The phase angle will be almost in phase with the applied voltage and Z will be approximately equal to R . In such a circuit the voltage drop across R is almost equal to the applied voltage with negligible drop across the X_C .

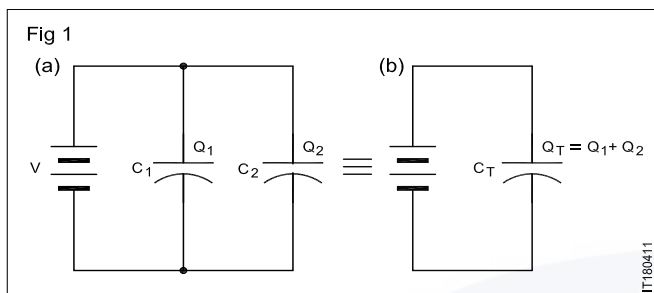
Capacitors in series & parallel

Objectives: At the end of this lesson you shall be able to

- state the effect of capacitors in parallel and series in respect of total capacitance and reactance
- explain the use of series capacitors as voltage dividers
- explain why no power is dissipated by a pure capacitor
- list the losses in a practical capacitor
- explain the terms 'Quality factor' and 'Dissipation factor'
- calculate the energy stored in a capacitor
- explain the application of capacitors in power supplies and uses of capacitors.

CAPACITORS IN PARALLEL

When two capacitors are connected in parallel as shown in Fig 1a, it is possible to represent these two capacitors by a single capacitor of equivalent value.



The charge that accumulates on each capacitor is given by,

$$Q_1 = C_1 \cdot V$$

and

$$Q_2 = C_2 \cdot V$$

If C_1 and C_2 are to be represented by one capacitor, say of value C_T , as shown in Fig 1b, then, C_T must store a total charge $Q_T = Q_1 + Q_2$.

Therefore,

$$Q_T = C_T \cdot V$$

$$C_T \cdot V = C_1 \cdot V + C_2 \cdot V$$

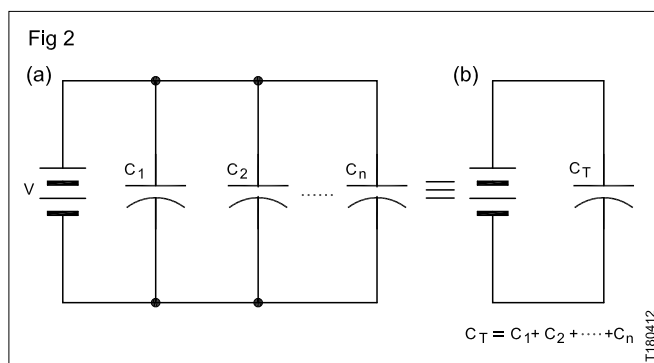
$$C_T \cdot V = V(C_1 + C_2)$$

or

$$C_T = C_1 + C_2$$

Generalizing, when several capacitors $C_1, C_2, \dots, C_n$, are connected in parallel, as shown in Fig 2, then the effective capacitance value C_T of this connection will be,

$$C_T = C_1 + C_2 + C_3 + \dots + C_n$$



Hence, when capacitors are connected in parallel, the effective capacitance increases.

It is interesting to note that, the equation for parallel-connected capacitors is similar to the series connection of resistors. The reason why an increase in capacitance results is due to the effective increase in the plate area, A (recall $C \propto A$).

When capacitors of different voltage ratings are connected in parallel, the maximum working voltage (DCWV) of the combination will be that of the lowest voltage rating capacitor amongst the group of capacitors.

Example: Three capacitors of specifications $15 \mu\text{F} - 25\text{V}$; $100 \mu\text{F} - 20\text{V}$; and $1 \mu\text{F} - 50\text{V}$ are connected in parallel. Determine,

- The total capacitance.
- The maximum voltage that can be applied to the circuit.
- The maximum charge that the parallel-connected group can store without damaging any capacitor in the connection.

SOLUTION

a. $C_T = C_1 + C_2 + C_3 = 15 \mu\text{F} + 100 \mu\text{F} + 1 \mu\text{F}$
 $= 116 \mu\text{F}$

- b. Maximum working voltage of the combination is the minimum working voltage of the capacitor in the group. In the group, $100 \mu\text{F}$ has the minimum DCWV of 20V .

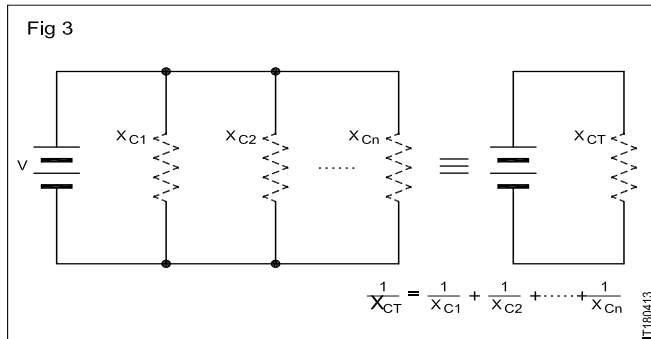
Hence, the maximum DCWV of the parallel combination is 20 volts only.

c. $Q_T = C_T \cdot V = 116 \mu\text{F} \times 20 \text{ Volts} = 2320 \mu\text{C}$.

Effective capacitive reactance of capacitors in parallel

When capacitors are connected in parallel, the total capacitance C_T increases. This causes the overall reactance of the circuit to decrease because, $X_C = 1/2\pi fC$.

Fig 3 shows the capacitive reactance of individual capacitors connected in parallel.



From Fig 3, it can be seen that X_C 's in parallel is similar to the resistances in parallel. Therefore, the total capacitive reactance X_{CT} of capacitors in parallel is given by,

$$\frac{1}{X_{CT}} = \frac{1}{X_{C1}} + \frac{1}{X_{C2}} + \frac{1}{X_{C3}} + \dots + \frac{1}{X_{Cn}}$$

Example: Two capacitors of $8\mu\text{F}$ -400V and $4\mu\text{F}$ -400V are connected in parallel across a 240 V, 50 Hz supply. Calculate,

- the total capacitive reactance
- the total current drawn from the supply.

SOLUTION

$$\begin{aligned} \text{a. } X_{C1} &= 1/2\pi \times 3.141 \times 50 \times 8 \times 10^{-6} = 398 \Omega \\ X_{C2} &= 1/1 \times 3.141 \times 50 \times 4 \times 10^{-6} = 796 \Omega \end{aligned}$$

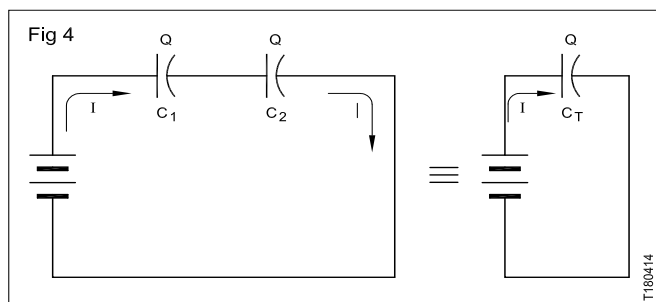
$$\frac{1}{X_{CT}} = \frac{1}{X_{C1}} + \frac{1}{X_{C2}}$$

$$\text{or } X_{CT} = \frac{X_{C1} \times X_{C2}}{X_{C1} + X_{C2}} = \frac{398 \Omega \times 796 \Omega}{398 \Omega + 796 \Omega} \approx 256 \Omega$$

$$\text{b. } I_T = \frac{V}{X_{CT}} = \frac{240\text{V}}{256 \Omega} \approx 0.90 \text{ A.}$$

CAPACITORS IN SERIES

When two capacitors are connected in series as shown in Fig 4, the charging current is the same throughout the circuit as it is in a series circuit.



Since the current is the same, the same quantity of charge, say Q is deposited on each capacitor. The voltage across any capacitor is given by the formula,

$$Q = CV \text{ or } V = \frac{Q}{C}$$

Hence, the voltage to which each capacitor charges is given by,

$$V_1 = \frac{Q}{C_1} \text{ and } V_2 = \frac{Q}{C_2}$$

But, by Kirchhoff's voltage law, $V_s = V_1 + V_2$.

$$\text{Therefore, } V_s = \frac{Q}{C_1} + \frac{Q}{C_2}$$

$$\text{or } V_s = Q \left(\frac{1}{C_1} + \frac{1}{C_2} \right)$$

If C_T was the equivalent capacitor of the series connection as shown in Fig 4, for C_T to charge to voltage V is given by,

$$V = \frac{Q}{C_T}$$

Therefore,

$$\frac{Q}{C_T} = V = \left(\frac{1}{C_1} + \frac{1}{C_2} \right) \text{ or } \frac{1}{C_T} = \frac{1}{C_1} + \frac{1}{C_2}$$

$$\text{or } C_T = \frac{C_1 \times C_2}{C_1 + C_2}$$

Generalizing, when n capacitors are connected in series, the effective capacitance of the series connection is given by,

$$\frac{1}{C_T} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \dots + \frac{1}{C_n}$$

Note that this equation of capacitors in series is similar to a parallel connection of resistors. Therefore, the effective capacitance of capacitors in series results in a decrease in the capacitance with the total capacitance always less than the smallest value of the capacitor in the group. The reason for this may be thought of as resulting from an effective increase in distance d , of the capacitor plates (recall $C \propto 1/d$).

Tip: When n capacitors each of same value C are connected in series, the total capacitance of such a series connection is given by the simple formula,

$$C_T = \frac{C}{n}$$

Example: Two capacitors $5\mu\text{F}$ - 20V and $10\mu\text{F}$ - 20 V are connected in series across a 30-V dc supply. Calculate,

- the total capacitance of the series combination
- the charge on each capacitor
- the voltage across each capacitor.

SOLUTION

$$a. C_T = \frac{C_1 C_2}{C_1 + C_2} = \frac{5 \times 10}{5 + 10} = \frac{50}{15} = 3 \frac{1}{3} \mu F$$

b. Since it is a series circuit, current is the same throughout the circuit. Hence, $Q_T = Q_1 = Q_2 = Q$

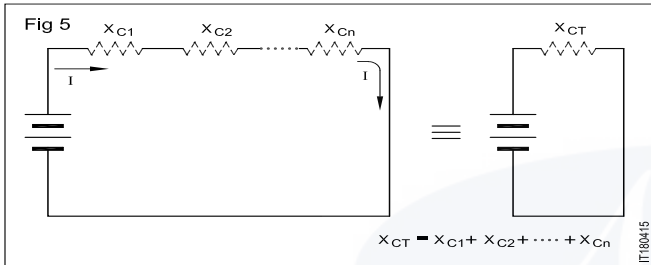
$$\text{Therefore, } Q = C_T \cdot V = 3 \frac{1}{3} \mu F \times 30 V = 100 \mu C.$$

$$c. V_2 = \frac{Q}{C_2} = \frac{100 \mu C}{5 \mu F} = 20 V$$

Tip: In a series connection, the capacitor with a smaller capacitance value will have a higher voltage drop across it.

Effective capacitive reactance of capacitors in series

Fig 5 shows the capacitive reactance of individual capacitors connected in series. This is similar to resistors connected in series.



Hence, as shown in Fig 5b, the total capacitive reactance will be,

$$X_{CT} = X_{C1} + X_{C2} + X_{C3} + \dots + X_{Cn}$$

The current through capacitors in series can be calculated as,

$$I = \frac{V}{X_{CT}}$$

Voltage across each capacitor can be found using the formula,

$$V_C = I \cdot X_C$$

Example: Two capacitors $8 \mu F$, $400V$ and $4 \mu F$, $400V$ are connected in series across mains supply of $240V$, 50-Hz . Calculate,

- the total capacitive reactance
- the current drawn by the capacitors
- the voltage across each capacitor.

SOLUTION

$$a. X_{C1} = \frac{1}{2\pi f C_1} = \frac{1}{2\pi \times 50 \text{ Hz} \times 8 \times 10^{-6} \text{ F}} = 397 \Omega$$

$$X_{C2} = \frac{1}{2\pi f C_2} = \frac{1}{2\pi \times 50 \text{ Hz} \times 4 \times 10^{-6} \text{ F}} = 796 \Omega$$

$$X_{CT} = X_{C1} + X_{C2} = 397 \Omega + 796 \Omega = 1193 \Omega.$$

$$b. I = \frac{V}{X_{CT}} = \frac{240 V}{1193 \Omega} = 0.20 \text{ ampere.}$$

$$c. V_{C1} = I X_{C1} = 0.20 A \times 397 \Omega = 80 V.$$

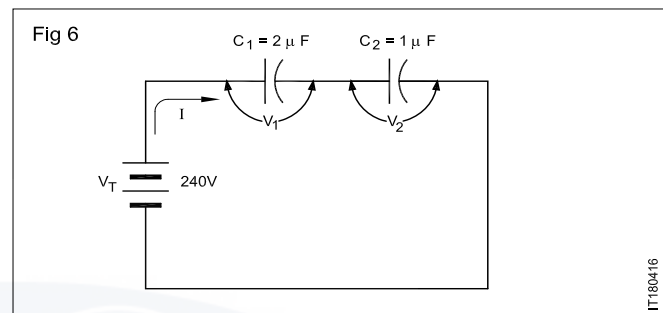
$$V_{C2} = I X_{C2} = 0.20 A \times 796 \Omega = 160 V.$$

Note that the total applied voltage divides in such a way, that higher voltage $160V$, appears across the smallest value capacitor ($4 \mu F$) of the group.

Series capacitors as voltage dividers

When two capacitors of different values and different working voltages are connected in series as shown in Fig 6, the voltage across the capacitors is given by,

$$V_1 = \frac{Q}{C_1} \text{ and } V_2 = \frac{Q}{C_2}.$$



Since the capacitors are in series, current through the capacitors is the same and, hence, the charge Q across them will be the same.

$$\text{Therefore, } Q = C_1 V_1 = C_2 V_2.$$

From this equation it can be seen that, since Q is same, the smaller capacitor will have the larger voltage across it, and vice versa.

In general, the following simple formula can be used to find the voltage across capacitors connected in series;

$$V_C = \frac{C_T}{C} \times V_T$$

where,

V_C is the voltage across any chosen capacitor

C_T is the total capacitance of the series connection

C is the capacitance of the chosen capacitor

and V_T is the total applied voltage.

A detailed method of calculating voltage across series capacitors as discussed earlier is by finding Q and this can also be used.

For the circuit shown in Fig 6, using the formula,

$$V_C = \frac{C_T}{C} \times V_T$$

voltage across capacitors is

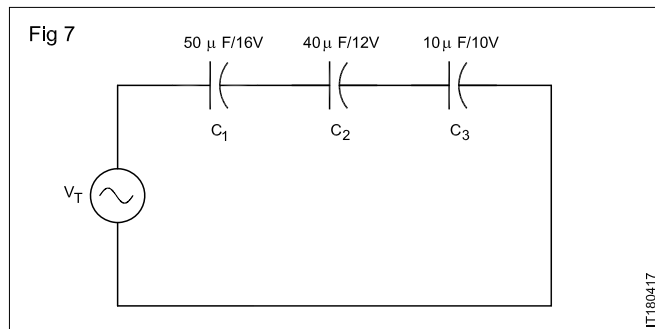
$$V_1 = \frac{C_T}{C_1} \times 240 = \frac{2/3}{2} \times 240 = 80 \text{ volts}$$

$$V_2 = \frac{C_T}{C_2} \times 240 = \frac{2/3}{1} \times 240 = 160 \text{ volts}$$

This method is applicable to series capacitances as voltage dividers for either dc or ac voltage, as long as there is no series resistor.

Example: Three capacitors $50 \mu\text{F}/16\text{V}$, $40 \mu\text{F}/12\text{V}$ and $10 \mu\text{F}/10\text{V}$ are connected in series as shown in Fig 7. Calculate,

- the maximum working voltage of this combination.
- the voltage across each capacitor.



SOLUTION

- Since C_3 is the smallest capacitor, it will have the highest voltage across it. This drop must not exceed its voltage rating, which must be limited to 10V.

Therefore, the maximum working voltage of the circuit, V_T can be found, if the voltage drop across C_3 (V_3) is limited to 10 volts.

Using the formula,

$$V_3 = \frac{C_T}{C_3} \times V_T$$

$$\text{maximum voltage } V_T = \frac{V_3 \times C_3}{C_T}$$

$$C_T = \frac{C_1 \cdot C_2 \cdot C_3}{C_1 \cdot C_2 + C_1 \cdot C_3 + C_2 \cdot C_3} = 6.9 \mu\text{F}$$

$$\text{Therefore, } V_T = \frac{10\text{V} \times 10 \mu\text{F}}{6.9 \mu\text{F}} = 14.5 \text{ volts.}$$

$$\text{Hence, } V_1 = \frac{C_T}{C_1} \times V_T = \frac{6.9 \mu\text{F}}{50 \mu\text{F}} \times 14.5\text{V} = 2 \text{ volts}$$

$$V_2 = \frac{C_T}{C_2} \times V_T = \frac{6.9 \mu\text{F}}{40 \mu\text{F}} \times 14.5\text{V} = 2.5 \text{ volts}$$

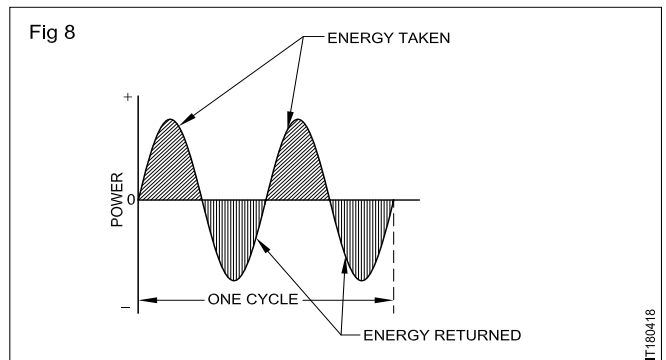
$$V_3 = \frac{C_T}{C_3} \times V_T = \frac{6.9 \mu\text{F}}{10 \mu\text{F}} \times 14.5\text{V} = 10 \text{ volts}$$

$$V_T = V_1 + V_2 + V_3 = 2 + 2.5 + 10 = 14.5 \text{ V.}$$

Power consumed in capacitor circuits

Capacitors merely store energy and return it while discharging. This means, there is no power dissipated by a capacitor unlike a resistor. Fig 8 shows the power or energy taken from the source as positive power, and power or energy returned to the source as negative energy.

In Fig 8, the negative power exactly cancels the positive power. Therefore, we can conclude two things about a circuit that contains only pure capacitance.



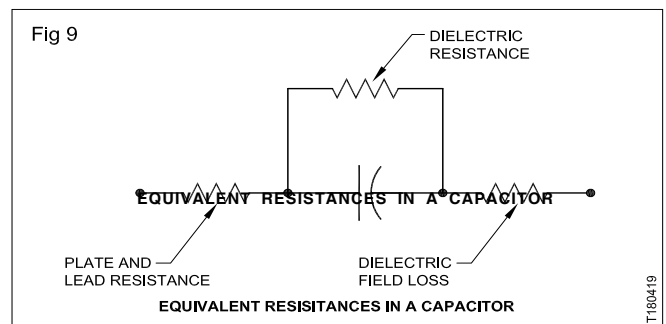
- 1 Current and voltage are 90° out of phase
- 2 The circuit uses no net energy or power.

But an ideal capacitor having only 'capacitance' and zero resistance cannot be practically made. As shown in Fig 9, due to their inherent resistances, capacitors do lose or consume some energy.

Referring to Fig 9, energy losses in capacitors can be classified as:

- 1 Resistance losses
- 2 Dielectric losses.

Resistance loss is caused primarily by current flowing through the resistance of the capacitor leads and plates and a very small current flowing through the high resistance of the dielectric material. The dielectric resistance appears as a parallel resistance as shown in Fig 9.



Dielectric loss results from continually reversing the polarity of the electric field through the dielectric when working on AC voltage. Each time the polarity changes, the molecular structure of the dielectric has to adjust to a new polarity. In this process, a small amount of electric energy is converted into heat energy. This dielectric loss appears as a small resistance in series with the capacitor.

The effects of the resistance losses and dielectric losses are often combined as the equivalent series resistance, referred to as ESR. A capacitor with a small ESR has a small energy loss.

The relative amount of energy lost in a capacitor can be indicated by any one of the following three terms.

- 1 Quality factor, Q.
- 2 Dissipation factor, DF
- 3 Power factor, PF

All these terms give the same information about a capacitor.

Quality factor, Q: Capacitors and inductors are built to provide reactance so that current can be controlled without converting electric energy into heat energy. However, it is impossible to build capacitors and inductors, especially inductors having zero resistance. *Quality* is the term used to rate a capacitor (or inductor) about its ability to produce reactance with as little resistance as possible.

Quality factor Q, can be defined as the ratio of reactance to resistance.

$$Q = \frac{\text{Reactance}}{\text{Resistance}}$$

For capacitors,

$$Q = \frac{\text{Capacitive reactance}}{\text{Resistance}} = \frac{X_C}{R}$$

where,

Q = quality factor of the capacitor at any given frequency ($Q \propto X_C$ and $X_C \propto 1/C$)

X_C = capacitive reactance at any given frequency

R = internal resistance of the capacitor (ESR)

Since both X_C and R have units of ohms, Q has no units.

Because X_C changes with frequency, Q is frequency-dependent.

Hence, the value of Q is always given for a specified frequency. If a capacitor has very little energy loss, R is low and Q is high of the order of more than 100.

Symbol Q used for representing quality factor shall not be confused with the symbol for charge which is also Q. Symbol Q for quality factor is retained as most books/technical brouchers use the symbol Q for quality factor.

Dissipation Factor, DF is the reciprocal of Quality factor. In other words, dissipation factor is the ratio of resistance to reactance. Hence, if Q is large, dissipation factor DF will be small number. If DF is given, Q can be found as $Q = 1/DF$ and vice-versa.

Power Factor, PF is the ratio of true power (P) to apparent power (P_{app}). In a capacitor, true power represents the energy lost per second, and apparent power represents the energy stored per second. If a capacitor has very little energy loss, its power factor is low, of the order of less than 0.01 or 1 percent.

For all capacitors, except for the lowest-quality capacitors, power factor PF and dissipation factor DF are essentially equal to each other. Summarizing, remember that capacitors with small energy losses have a high Q and low dissipation and power factors.

In low quality capacitors, such as electrolytic capacitors, energy losses can be considerable, of the order of 1 to 10 percent of the stored energy. However, with high quality non-electrolytic capacitors the losses are so small that they can be ignored for most applications.

Energy stored in a capacitor

Once a capacitor is charged, the voltage across it remains even when the source is removed. In fact, a capacitor is the only device that can store electrical energy apart from an electric voltaic cell (an inductor also stores energy but only as long as current is flowing through it). In some applications, capacitors can store a charge for even months. This property is used to indicate the measurements of pressure or temperature in remote locations. The energy stored in a capacitor is given by the equation

$$W = \frac{1}{2} CV^2 \text{ joules (J)}$$

where,

W is the energy stored in the capacitor in joules, J

C is the capacitance in farads, F

V is the final voltage across the capacitor in volts, V.

Example: Two capacitors 40 μ F/450V and 20 μ F/500 V are connected in parallel across a 400-V dc supply. Calculate the

- a) energy stored by each capacitor and b) the total energy stored.

SOLUTION:

For the 40 μ F,

$$W_1 = \frac{1}{2} C_1 V^2 = \frac{1}{2} \times 40 \times 10^{-6} \text{ F} \times (400 \text{ V})^2 = 3.2 \text{ J.}$$

For the 20 μ F,

$$W_2 = \frac{1}{2} C_2 V^2 = \frac{1}{2} \times 20 \times 10^{-6} \text{ F} \times (400 \text{ V})^2 = 1.6 \text{ J.}$$

$$\text{b) Total energy} = W_1 + W_2 = 3.2 \text{ J} + 1.6 \text{ J} = 4.8 \text{ J}$$

$$\text{or } W_T = \frac{1}{2} C_T V^2 = \frac{1}{2} \times 60 \times 10^{-6} \text{ F} \times (400 \text{ V})^2 = 4.8 \text{ J.}$$

Capacitors charged to the energy levels as in the example are dangerous. To avoid shock/prevent damage to measuring instruments discharge them before handling.

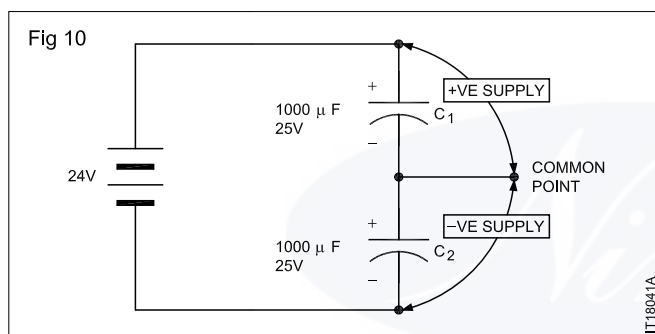
The expression for stored energy in a capacitor ($CV^2/2$) is similar to that for an inductor ($LI^2/2$). But unlike a capacitor, an inductor stores energy only so long as the current is maintained.

Use of capacitor to obtain dual supply

Most of the electronic circuits such as radio receivers, amplifier etc, require supply voltages such as +6V or +12V etc. Since these circuits use only + supplies, these are called single supply circuits.

On the other hand, some electronic circuits, require both +ve and -ve supplies such as +12V, -12V or +15, -15v etc. These are called dual supply circuits.

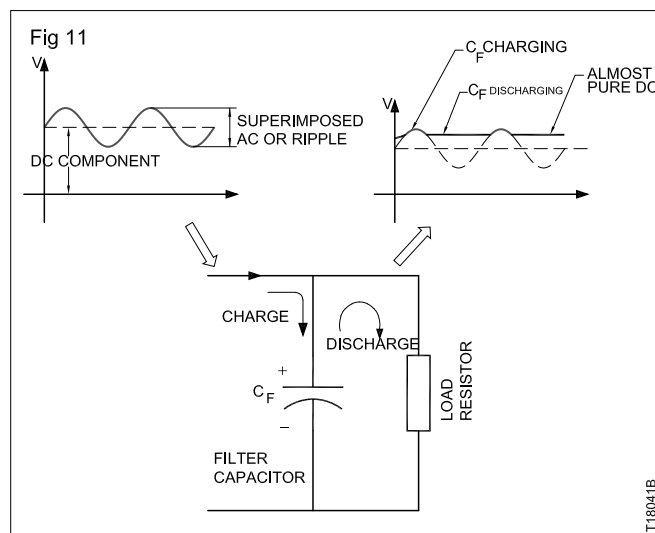
Most of the general power supplies provide only single supply. Capacitors can be used to convert an available single supply into a $\pm$ (Plus-minus) supply in a very simple way as shown in Fig 10.



Details of dual supplies are dealt with in further units.

Capacitors as ripple filters

The small unwanted AC voltage superimposed over a pure DC signal as shown in Fig 11 is known as ripple voltage. Such ripple voltage is most common in low quality *battery eliminators* used to obtain DC voltage from AC mains supply.

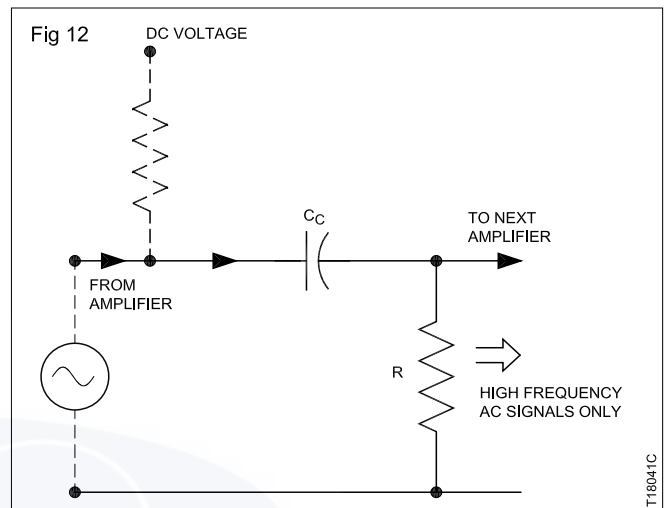


All electronic circuits invariably need pure DC voltage. If the ripple in the DC is large, the electronic circuits do not behave as expected. For example, a ripple in the DC voltage given to a radio receiver produces an unwanted humming noise in the speaker.

The property of charging and discharging of capacitors is made use of to remove or to filter off the unwanted ripple as shown in Fig 11.

Capacitors for AF and RF coupling

One of the main uses of capacitors is in coupling the signal between the two stages of amplifiers as shown in Fig 12.



Capacitors are used for coupling because of the following reasons.

- 1) It provides low reactance at high frequencies, resulting in low AC voltage drop across C.
- 2) The entire dc voltage is dropped across C with none across R.

As a result, the voltage across R is only of the desired high frequency AC signal and not DC current or very low frequency signal. This application of the capacitor C_C , therefore, is called ac coupling.

The value of C_C to be used depends upon the lower limit of the frequency to be coupled. The table given below gives an idea of the value of capacitor to be used for different frequency ranges.

Frequency range	Lower frequency limit chosen	Value of C_C used
Low audio frequencies	100 Hz	1 μ F
Audio frequencies (mid)	1000 Hz	0.1 μ F
Audio frequencies (high)	10 kHz	0.01 μ F
Radio frequencies	1000 Hz	100 pF
Very high frequencies	100 MHz	1 pF

Further details on coupling capacitors is dealt in the unit on amplifiers.

Series and parallel resonance

Objectives: At the end of this lesson you shall be able to

- state the meaning of audio frequency and radio frequency amplifiers
- state the meaning of broad-band narrow-band and tuned amplifiers
- state the meaning of resonance in LC circuits
- state the condition for resonance in a series and parallel LC circuit
- list the characteristics of LC circuits at resonance
- explain the important terms used in LC circuits & relate them
- explain the storage action in parallel LC circuit
- explain the terms make up current and tank circuit
- list a few applications of series and parallel resonance circuits.

TUNED Amplifiers

The amplifiers discussed so far amplified audio amplifiers and hence are referred to as audio amplifiers. These audio amplifiers amplified the wide range of signal frequencies (20Hz to 20 KHz) in the audio range. Such amplifiers which amplify a wide band of frequencies or a broad band of frequencies are referred to as *wide band amplifiers* or *broad band amplifiers*.

Often, amplifiers may be required to amplify a very narrow band of frequencies. Such amplifiers which are designed to amplify a single frequency or a very narrow band of frequencies are referred to as *narrow-band amplifiers*.

Narrow-band amplifiers are commonly used in higher frequency ranges. One such application of narrow-band amplifiers is radio and television receivers where, a narrow band of *radio frequency (RF)* are amplified by amplifiers.

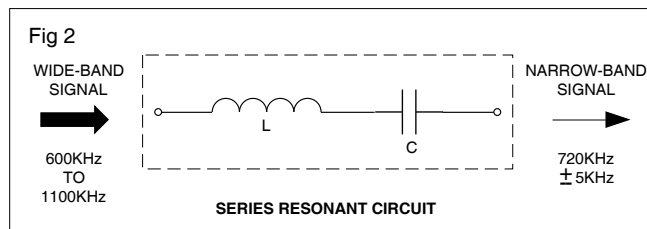
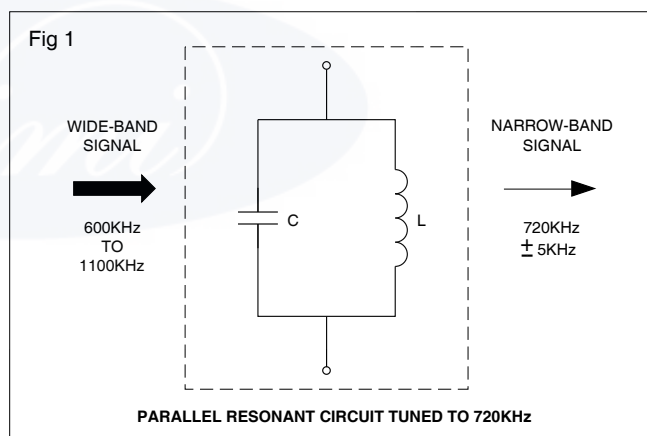
In a radio broad casting station, the audio signal (voice or music) is superimposed on a high frequency (RF) much above the audible frequency range. This process of raising the audio signal frequencies to radio frequencies (RF) is called *modulation* (details discussed in further lessons). Some of the radio frequencies over which audio signals are sent (broad-cast) by radio stations are listed below;

Station	Transmitted Frequency
CHENNAI - A	720 KHz
CHENNAI - Vividbharati	783 KHz

When the RF signal (carrying audio signal) reaches the receiving antenna/aerial, the signal will be very weak and, hence, needs to be amplified before extracting the audio information from it. To amplify this weak RF signal, the amplifier used should be capable of amplifying a very narrow band of frequencies around that single frequency (either 720 KHz or 783 KHz) such that only the desired station, say Chennai-A station or Chennai-B station only is received. This is how a radio receiver can separate one station from another. Amplifiers which are designed to amplify high radio frequencies are referred to as *RF amplifiers*. Amplifiers used to amplify such very narrow band radio frequencies are called *narrow band RF amplifiers*.

or *tuned amplifiers*. They are called tuned amplifiers because, the amplifier is tuned to amplify a specific frequency only. To obtain this tuning, these amplifiers use a tuned circuit. A tuned circuit generally consists of an inductor (L) and a capacitor (C) as shown in Fig 1 and Fig 2.

The speciality of this LC network is its property, known as *resonance*. Whenever an inductor and a capacitor are connected in parallel as shown in Fig 1 or in series as shown in Fig 2, the circuits exhibit a special phenomenon known as *resonance*. Resonance, in simple terms, means, the ability to select a particular frequency from the fed signal consisting of signals of different frequencies (wide band of frequencies).



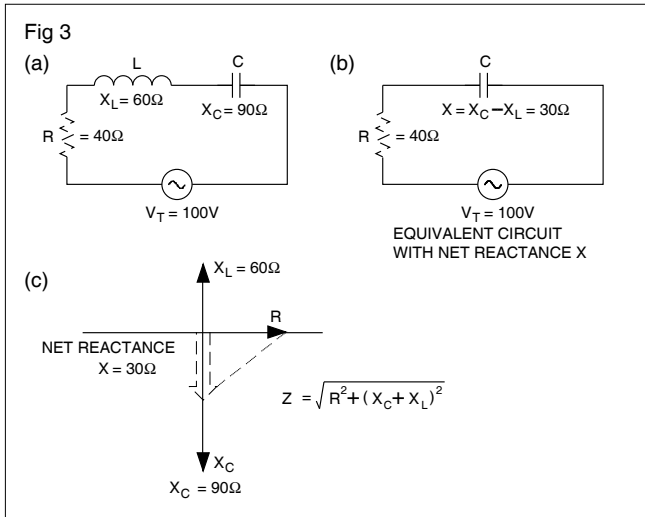
SERIES RESONANCE CIRCUIT

Impedance of series resonance circuit

A simple series LC circuit shown in Fig 3a. In this series LC circuit,

- resistance R is the total resistance of the series circuit (internal resistance) in ohms,
- X_L is the inductive reactance in ohms, and
- X_C is the total capacitive reactance in ohms.

In the circuit at Fig 3, since the capacitive reactance(90W) is larger than inductive reactance(60W), the net reactance of the circuit will be capacitive. This is shown in Fig 3b.



If the capacitive reactance was smaller than inductive reactance the net reactance of the circuit would have been inductive.

All though the unit of measure of reactance and resistance is the same(ohms) the impedance, Z of the circuit is not given by the simple addition of R, X_L and X_C . This is because, X_L is $+90^\circ$ out of phase with R and X_C is -90° out of phase with R.

Hence the impedance Z of the circuit is the phasor addition of the resistive and reactive components as shown by dotted lines in Fig 3c. Therefore, Impedance Z of the circuit is given by,

$$Z = \sqrt{R^2 + (X_C - X_L)^2}$$

If X_L were greater than X_C , then the absolute value of impedance Z is will be,

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

For the circuit in Fig 3(a), total impedance Z is,

$$Z = \sqrt{R^2 + (X_C - X_L)^2}$$

$$Z = \sqrt{40^2 + 30^2}$$

$$Z = 50\Omega, \text{ Capacitive (because } X_C > X_L \text{)}$$

Current I through the circuit is given by,

$$I = \frac{V}{Z} = \frac{100}{50} = 2 \text{ Amps.}$$

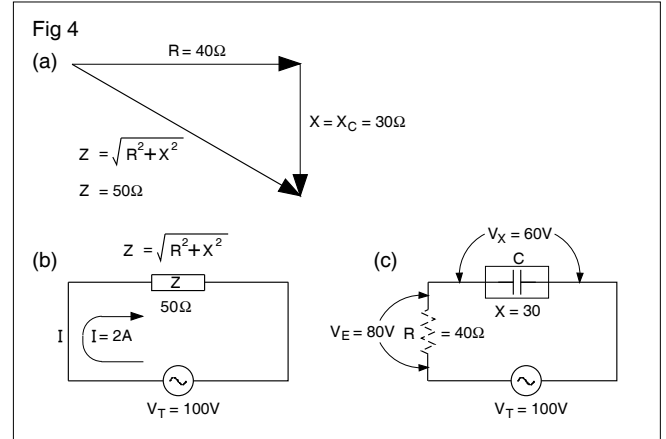
Therefore, the voltage drop across the components will be,

$$V_R = \text{voltage drop across } R = I \cdot R = 2 \times 40 = 80 \text{ volts}$$

$$V_L = \text{voltage drop across } L = I \cdot X_L = 2 \times 60 = 120 \text{ volts}$$

$$V_C = \text{voltage drop across } C = I \cdot X_C = 2 \times 90 = 180 \text{ volts.}$$

Since V_L and V_C are of opposite polarity, the net reactive voltage V_X is $= 180 - 120 = 60V$ as shown in Fig 4.



Note that the applied voltage is not equal to the sum of voltage drops across reactive component X and resistive component. This is again because the voltage drops are not in phase. But the phasor sum of V_R and V_X will be equal to the applied voltage as given below,

$$\begin{aligned} V_T &= \sqrt{V_R^2 + V_X^2} \\ &= \sqrt{V_R^2 + (V_L - V_C)^2} \\ &= \sqrt{80^2 + 60^2} = 100 \text{ volts (applied voltage).} \end{aligned}$$

Phase angle θ of the circuit is given by,

$$\theta = \tan^{-1} \frac{X_C - X_L}{R}$$

Condition at which current through the RLC Series circuit is maximum

From the formula,

$$Z = \sqrt{R^2 + (X_C - X_L)^2} \quad \text{it is clear that the total impedance Z of the circuit will become purely resistive when,}$$

$$\text{reactance } X_L = X_C$$

In this condition, the impedance Z of the circuit will not only be purely resistive but also minimum.

Since the reactance of L and C are frequency dependent, at some particular frequency say f_r , the inductive reactance X_L becomes equal to the capacitive reactance X_C .

In such a case, since the impedance of the circuit will be purely resistive and minimum, current through the circuit will be maximum and will be equal to the applied voltage divided by the resistance R.

Series resonance

From the above discussions it is found that in a series RLC circuit,

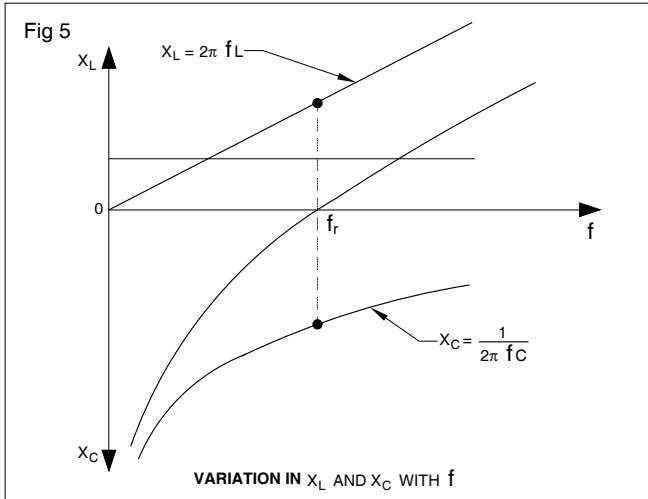
$$\text{Impedance } Z = \sqrt{R^2 + (X_L - X_C)^2}$$

$$\text{Current } I = \frac{V}{Z}$$

and,

$$\text{Phase angle } \theta = \tan^{-1} \frac{X_L - X_C}{R}$$

If the frequency of the signal fed to such a series LC circuit (Fig 3a) is increased from 0 Hz, as the frequency is increased, the inductive reactance ($X_L = 2\pi fL$) increases linearly and the capacitive reactance ($X_C = 1/2\pi fL$) decreases exponentially as shown in Fig 5.



As shown in Fig 5, at a particular frequency called the **resonance frequency**, f_r , the sum of X_L and X_C becomes zero ($X_L - X_C = 0$).

From Fig 5 above, at resonant frequency f_r ,

- Net reactance, $X = 0$ (i.e. $X_L = X_C$)
- Impedance of the circuit is minimum, purely resistive and is equal to R
- Current I through the circuit is maximum and equal to V/R
- Circuit current, I is in-phase with the applied voltage V (i.e. Phase angle $\theta = 0$).

At this particular frequency f_r called resonance frequency, the series RLC is said to be in a condition of **series resonance**.

Resonance occurs at that frequency when,

$$X_L = X_C \text{ or } 2\pi fL = 1/2\pi fC$$

Therefore, **Resonance frequency**, f_r is given by,

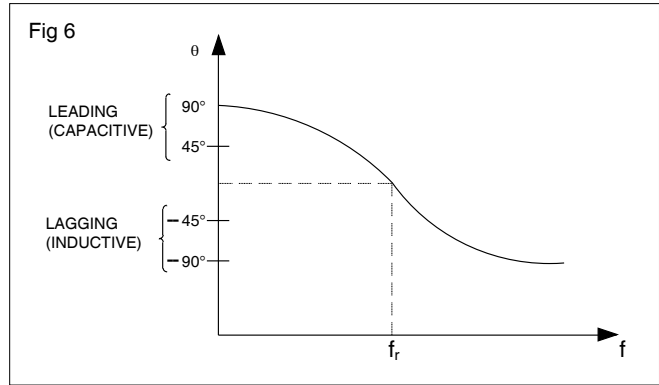
$$f_r = \frac{1}{2\pi\sqrt{LC}} \text{ Hz} \quad \dots[1]$$

Reactance of series RLC above and below resonance frequency f_r

Fig 6 shows the variation of net reactance of a RLC circuit with the variation in frequency.

From Fig 6 above, it can be seen that the,

- net reactance is zero at resonant frequency f_r
- net reactance is capacitive below the resonant frequency f_r



- net reactance is inductive above the resonant frequency f_r .

Selectivity or Q factor of a series RLC circuit

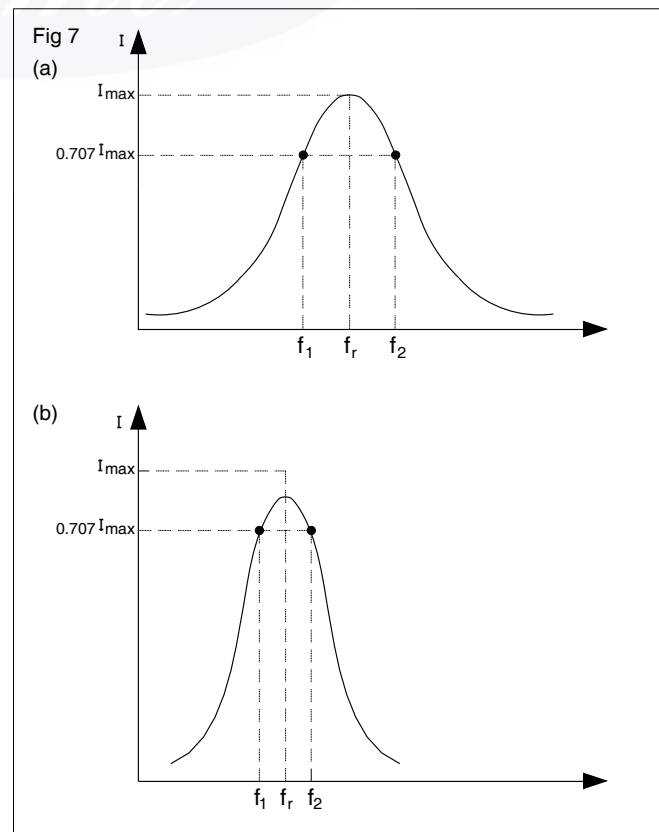
Figs 7a and 7b two graphs showing the current through series two different RLC circuits for frequencies above and below f_r . f_1 and f_2 are frequencies at which the circuit current is 0.707 times the maximum current, $I_{\max}$ or the -3dB points.

Fig 7 indicates that series RLC circuits select a band of frequencies around the resonant frequency, f_r . This band (f_1 to f_2) is called the **band width f** of the series RLC circuit.

$$\text{Bandwidth} = \Delta f = f_2 - f_1 \text{ Hz.}$$

where, f_2 is called the *upper cut off frequency* and f_1 is called the *lower cut off frequency* of the resonant circuit.

Comparing Figs 7a and 7b, it is seen that the bandwidth of 7b is smaller than that of 7a. This is referred to as the **selectivity** or **quality factor, Q** of the resonance circuit.



The RLC circuit having the response shown in Fig 7b is more selective than that of Fig 7a. The quality factor, Q of a resonance circuit is given by,

$$\text{Quality factor} = Q = \frac{f_r}{\Delta f} = \frac{f_r}{f_2 - f_1} \quad \dots[2]$$

If Q is very large, the bandwidth f will be very narrow and vice-versa. The Q factor of the series resonance circuit depends largely upon the Q factor of the coil (inductance) used in the RLC circuit.

Therefore,

$$Q \text{ of coil} = \frac{X_L}{R} = \frac{2\pi fL}{R}$$

since,

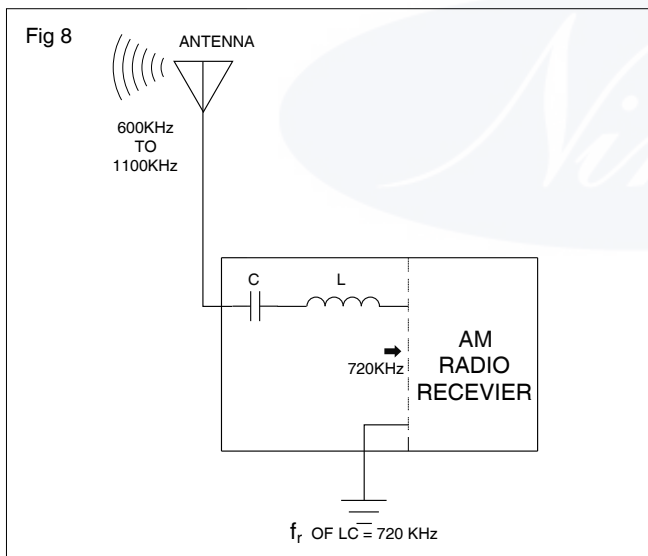
$$f_r = \frac{1}{2\pi\sqrt{LC}}$$

Q of the series RLC circuit is given by,

$$Q = \frac{1}{R} \cdot \frac{\sqrt{L}}{\sqrt{C}} \quad \dots[3]$$

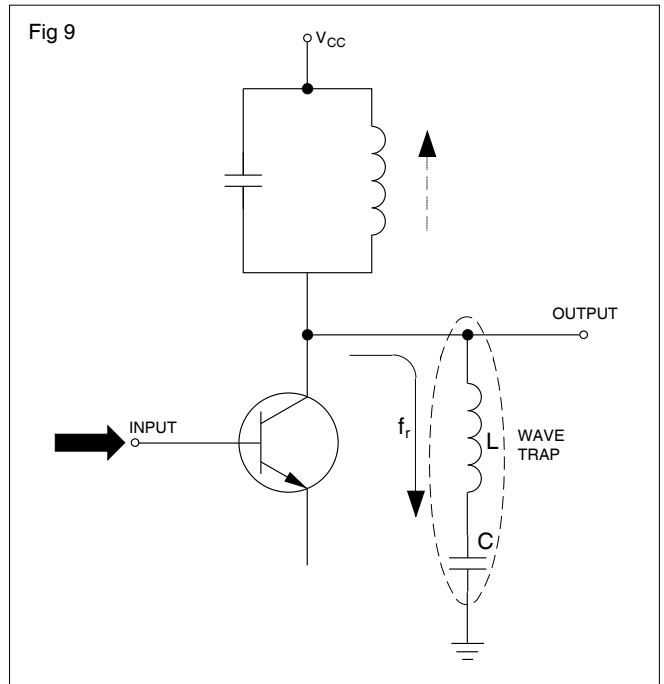
Application of series resonance circuits

A series resonance circuit can be used in any application where it is required to select a desired frequency. One such application is shown in Fig 8.



In Fig 8, the radio receiver antenna intercepts all the frequencies available in air. For instance, the antenna intercepts the frequencies transmitted by Chennai-A radio station, Chennai-Vividhabharthi station, and so on. The series LC circuit when tuned to 720KHz (f_r) will allow only the signal corresponding to Chennai-A radio station and rejects all other signal frequencies. This is how a radio can receive any one station although frequencies of many stations are in the air simultaneously.

A series resonance circuit can also be used to reject an undesired frequency (used as wave trap) as shown in Fig 9.



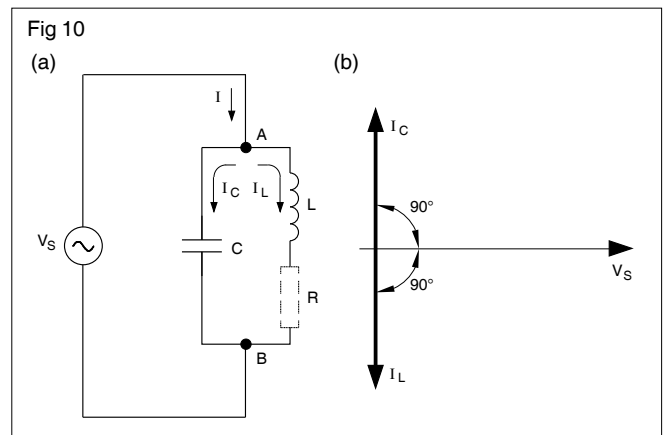
A *wave trap* is a resonant LC circuit tuned to the frequency to be rejected. Thus the output of the tuned amplifier in Fig 9 will not have the frequency for which the trap is tuned. This is because, at resonance the series LC of the wave trap, provides very low impedance. As the trap is connected across the collector and ground, the rejection frequency component is grounded.

Such wave traps are extensively used in very high and ultra-high frequency circuits such as television receivers, communication receivers etc.

PARALLEL RESONANCE

The circuit at Fig 10, having an inductor and a capacitor connected in parallel is called parallel LC circuit or parallel resonance circuit. The resistor R , shown in dotted lines indicate the internal dc resistance of the coil L . The value of R will be so small compared to the inductive reactance, that it can be neglected.

From Fig 10a, it can be seen that the voltage across L and C is same and is equal to the input voltage V_s .



By Kirchhoff's law, at junction A,

$$I = I_L + I_C.$$

The current through the inductance I_L (neglecting resistance R), lags V_S by 90° . The current through the capacitor I_C , leads the voltage V_S by 90° . Thus, as can be seen from the phasor diagram at Fig 10b, the two currents are out of phase with each other. Depending on their magnitudes, they cancel each other either completely or partially.

If $X_C < X_L$, then $I_C > I_L$, and the circuit acts capacitively.

If $X_L < X_C$, then $I_L > I_C$, and the circuit acts inductively.

If $X_L = X_C$, then $I_L = I_C$, and hence, the circuit acts as a purely resistive.

Zero current in the circuit means that the impedance of the parallel LC is infinite. This condition at which, for a particular frequency, f_r , the value of $X_C = X_L$, the parallel LC circuit is said to be in parallel resonance.

Summarizing, for a parallel resonant circuit, at resonance,

$$X_L = X_C,$$

$$Z_p = \infty$$

$$I_L = I_C$$

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

$$I = \frac{V}{Z_p} \approx 0$$

In a parallel resonance circuit, with a pure L (no resistance) and a pure C (loss-less), at resonance the impedance will be infinite. In practical circuits, however small, the inductor will have some resistance. Because of this, at resonance, the phasor sum of the branch currents will not be zero but will have a small value I . This small current I will be in phase with the applied voltage and the impedance of the circuit will be very high although not infinite.

Summarizing, the three main characteristics of parallel resonance circuit at resonance are,

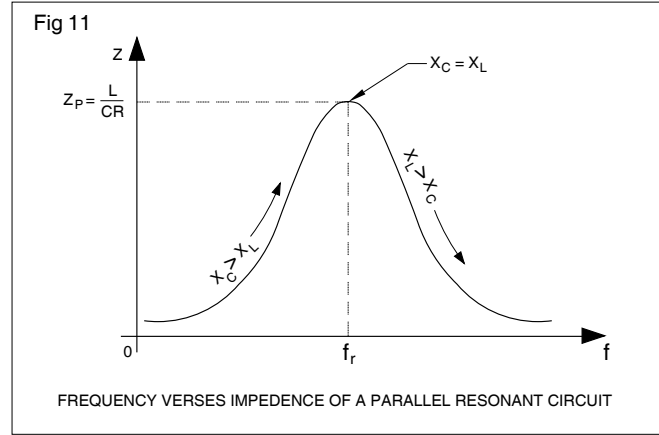
- phase difference between the circuit current and the applied voltage is zero
- maximum impedance
- minimum line current.

The variation of impedance of a parallel resonance circuit with frequency is shown in Fig 11.

In Fig 11, when the input signal frequency to the parallel resonance circuit is moved away from resonant frequency f_r , the impedance of the circuit decreases. At resonance the impedance Z_p is given by,

$$Z_p = \frac{L}{CR}$$

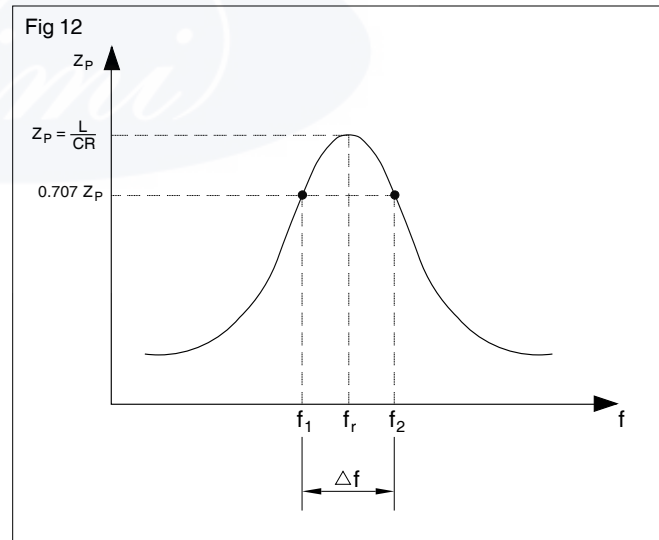
At resonance, although the circuit current is minimum, the magnitudes of I_L & I_C will be much greater than the line current. Hence, a parallel resonance circuit is also called *current magnification* circuit.



Bandwidth of parallel resonant circuits

As discussed in series resonance, all resonant circuits have the property of discriminating between the frequency at resonance (f_r), and those not at resonance. This discriminating property of the resonant circuit is expressed in terms of its **bandwidth(BW)**. In the case of series resonant circuits the response of the circuit at resonance frequency (f_r) is in terms of the line current (which is maximum), and in a parallel resonant circuit, it is in terms of the impedance (which is maximum).

The bandwidth of a parallel resonant circuit is also defined by the two points on either side of the resonant frequency at which the value of impedance Z_p drops to 0.707 or $1/\sqrt{2}$ of its maximum value at resonance, as shown Fig 12.



From Fig 12, the bandwidth of the parallel resonance circuit is,

$$\text{Bandwidth, BW} = \Delta f = f_2 - f_1$$

As can be seen in Fig 12, the value of Z_p is dependent on the resistance R of the coil ($Z_p = L/CR$). If R is less Z_p will be larger and vice versa. Since the bandwidth depends on Z_p and Z_p depends on R , we can say that the bandwidth of a resonant circuit depends upon the resistance associated with the coil. The resistance of the coil in turn decides the Q of the circuit. Thus, the Q of the coil decides the bandwidth of the resonant circuit and is expressed as,

$$\text{Bandwidth(BW)} = (f_2 - f_1) = \frac{f_r}{Q}$$

Storage action of parallel resonance circuit

At parallel resonance, though the circuit current is minimum(ideally zero), I_L and I_C will still be there. This I_L and I_C will be a circulating current in the closed loop formed by L and C. This circulating current will be very high at resonance. This circulating current flip-flops between the capacitor and inductor, alternately charging and discharging each. When a capacitor or an inductor is charged, it stores energy. When it is discharged it gives up the energy stored in it. The current inside the LC circuit switches the stored energy back and forth between L and C. If the inductor had no resistance and if the capacitor was loss-free, then, no more external energy would be required to retain this flip-flop or oscillation of charging and discharging. But, in a practical circuit, since ideal L and C cannot be obtained, some amount of the circulating energy is lost due to the resistance of the coil and the loss due to capacitor. This lost energy is the only energy the power supply source(V_s) must supply in the form of circuit current, I . This current, therefore, is called as **make-up current**. It is this storage action of the parallel-resonant circuit which gives rise to the term **tank circuit**, often used with parallel resonant circuits. Hence, parallel resonant circuits are also called tank circuits.

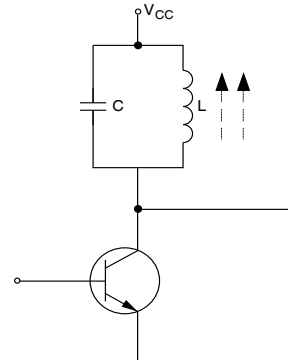
Table below gives a comparison between *series resonant* and *parallel resonant* circuit at frequencies above and below their resonant frequency f_r .

Property	Series circuit	Parallel circuit
	At resonant frequency	
Resonant frequency, f_r	$\frac{1}{2\pi\sqrt{LC}}$	$\frac{1}{2\pi\sqrt{LC}}$
Reactance	$X_L = X_C$	$X_L = X_C$
Impedance	Minimum ($Z_r = R$)	Maximum ($Z_r = L/CR$)
Current	Maximum	Minimum
Quality factor	$\frac{X_L}{R}$	$\frac{X_L}{R}$
Bandwidth	$\frac{f_r}{Q}$	$\frac{f_r}{Q}$
Above resonant frequency		
Reactance	$X_L > X_C$	$X_C > X_L$
Impedance	Increases	Decreases
Phase difference	The current lags behind the applied voltage.	The current leads the applied voltage.
Type of reactance	Inductive	Capacitive
Below resonant frequency		
Reactance	$X_C > X_L$	$X_L > X_C$
Impedance	Increases	Decreases
Phase difference	The current leads the applied voltage.	The current lags behind the applied voltage.
Type of reactance	Capacitive	Inductive

Application of parallel resonant circuits

Parallel resonance circuits or tank circuits are commonly used in almost all high frequency circuits. Tank circuits are used as collector load in class-C amplifiers instead of a resistor load as shown in Fig 13.

Fig 13



NOTE: Details of class-C amplifiers are discussed in further lessons.

Tank circuits are used in circuits known as oscillators which are designed to generate ac signals using dc supply.

NOTE: Oscillators are discussed in detail in further lessons.

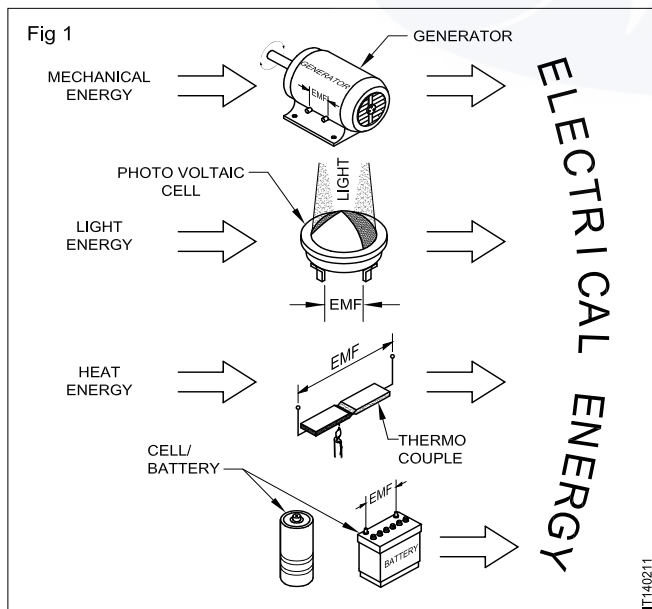
Types of cells & Batteries

Objectives: At the end of this lesson you shall be able to

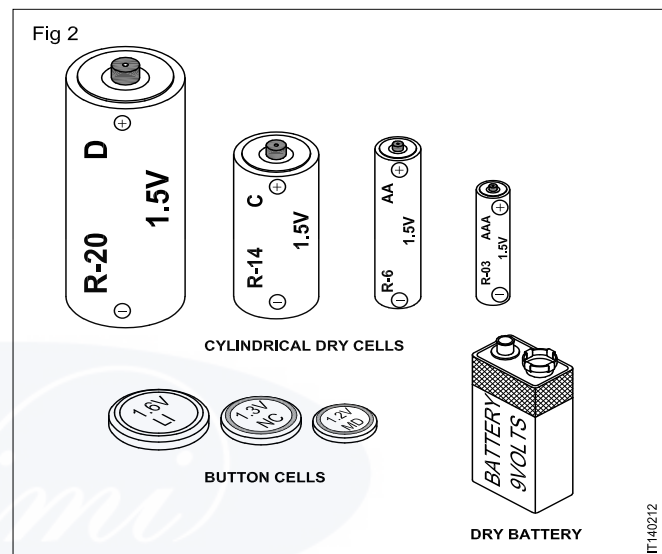
- state the difference between battery and other power sources and classifications
- name the basic unit of a battery and popular sizes
- state the difference between dry and wet cells, primary and secondary cells
- state the need for mechanical zero setting of the meters
- state the precautions to be taken while measuring the output of cells
- reason, why, not all types of batteries are suitable for all applications
- state the basic working principle of a multimeter and steps to be following while using multimeter
- state the method of identification and different scales of a multimeter.

POWER SOURCES

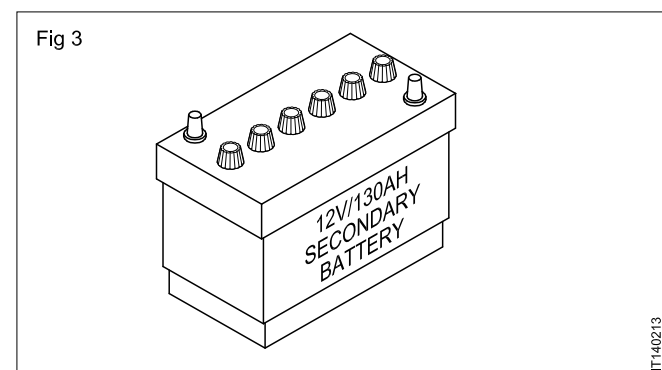
Devices that produce electricity are generally termed as Power sources. These power sources produce electricity by converting some form of energy into electrical energy. As shown in Fig 1, all power sources must first be supplied with external energy such as heat, light or mechanical energy before they can produce electricity with an exception in the case of cell/battery. Batteries are different from the other types of power sources because, energy is self-contained in batteries. Therefore, no energy need be supplied from outside for the battery to produce electricity. Hence batteries are one of the most important power sources. In a battery, electrical energy is produced by the chemicals contained within the battery. Cells are the basic units of a battery. Several cells forms to make a battery. Batteries are classified mainly under two categories



Primary Batteries - Converts chemical energy to electrical energy. This uses the chemicals within it to start the action of energy conversion. The most common types of primary cells and batteries are shown in Fig 2.



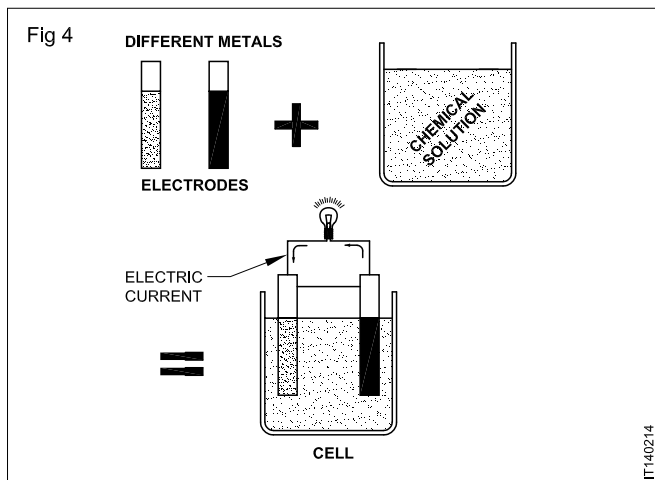
Secondary batteries - These batteries must be first charged with electrical energy. Once the battery is fully charged, it will then convert chemical energy to electrical energy. Secondary batteries first store electrical energy supplied to it and then supply electrical energy as and when required. Hence secondary batteries are commonly called **storage batteries**. A typical secondary storage battery is shown in Fig 3.



A battery may consist of two or more number of cells. The battery shown in Fig 3 has six cells of 2V each. These cells are connected in series to give 12V at battery terminals.

THE CELL

A cell consists of a pair of metal strips called **electrodes** and dipped in a chemical solution called **electrolyte** as shown in Fig 4.



Dry and Wet cells

The electrolyte can be in liquid form or a paste form. Cells with paste form electrolyte are known as **DRY** cells. Cells with liquid form of electrolyte are called **WET** cells.

Primary and Secondary cells

Primary cells are those which once fully used has to be thrown-out or destroyed. This is because the electrodes and electrolyte used in this type of cells cannot be reused. Hence, primary cells are *non-rechargeable* cells. Generally, the electrolyte used in primary cells is of paste form.

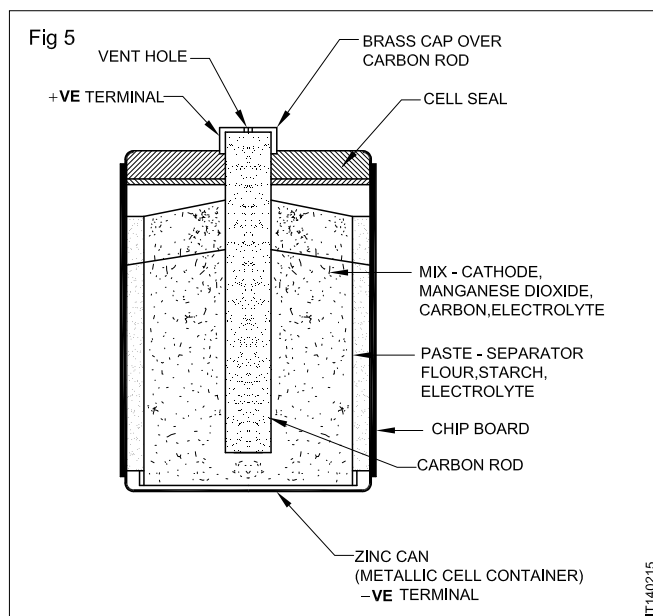
Secondary cells are those which once used can be reused by charging them. Hence, secondary cells are rechargeable cells. Generally, the electrolyte used in secondary cells is in liquid form. However, there are rechargeable cells with paste form electrolyte also.

In this lesson the commercial aspects of primary cells are discussed. Secondary cells are discussed in further lessons.

Dry cells and batteries

As the electrolyte used in dry cells is in paste form, it does not spill or leak. Hence, dry cells are used extensively in portable electrical and electronic gadgets. Typical constructional details of a zinc-carbon dry cell is shown in Fig 5.

The two electrodes of dry cells are brought out and are available as +ve and -ve terminals of the cell. Usually the metallic cell container serves as the -ve of the cell as shown in Fig 5. The voltage that appears across the terminals depends upon the electrodes and the chemicals used in the cell. The voltage of a cell is so made as to suit the commercial requirement. Generally the voltage across the terminals of a dry cell range between 1.2 to 1.5 volts.



Dry cells and batteries are available in several shapes and sizes to suit commercial requirements. Some popular shapes of dry cells were shown in Fig 2.

Technically, any particular type of cell is defined by the materials used as electrodes and electrolytes in that cell. A dry cell with **zinc** as the -ve electrode, **carbon** as the +ve electrode with **zinc chloride** as the electrolyte is referred to as **zinc-carbon** cell or **zinc chloride** cell.

Similarly a dry cell which uses an alkaline solution as electrolyte is called an **Alkaline cells**.

A Chart on *Types of cells/batteries* given at the end of this lesson lists some popular dry cells along with the names of the materials used for the +ve, -ve electrodes, the electrolyte used, the available sizes, the rated output voltage and their applications.

The use of different materials for their electrodes and electrolytes results in different voltage, current rating discharge characteristics and the shelf life (life of the battery if kept unused).

NOTE: Not all types of cells are suitable for all applications. This is because some appliances draw high initial current or current in pulses which may not suit the discharge characteristics of the cell.

Weak, dead cell

Dry cells are used in various gadgets like flash lights, tape recorders etc, the cells convert the chemical energy built into them into electrical energy. In doing so, the dry cell slowly gets consumed. This means, the voltage across the cell terminals decreases and the current it can supply to the connected load becomes less and less. A stage will reach when the dry cell is no more capable of supplying sufficient voltage/current through the connected load. Then the cell is said to have become **weak** or **dead**.

As a thumb rule, dry cell can be declared unfit for use if, the voltage across its terminals is less than 75% of its rated output voltage.

Example: A used zinc chloride dry cell with a rated voltage of 1.5 volts has 1.1 volts across its terminals. Find whether the cell is usable or not.

Rated output voltage of the cell is 1.5V.

Measured output of the cell is 1.1V.

% Measured output with respect to rated output is

$$\frac{1.1}{1.5} \times 100 = 73.3 \%$$

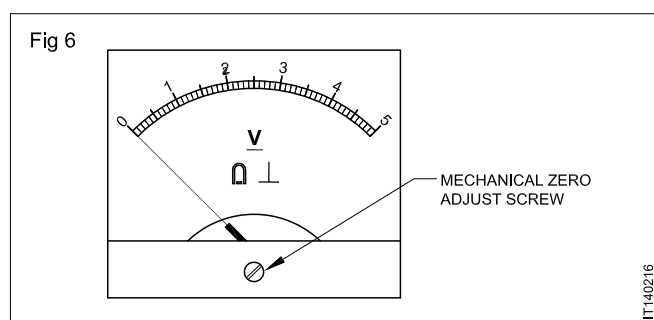
As the output is less than 75% of the rated output, the dry cell is unfit for use or can be treated as a **Weak cell**.

Measuring cell voltage

To measure the voltage of a dry cell, a voltmeter having a full scale deflection value greater than, but close to the rated output of the dry cell, should be used. Also the meter's minimum measurable voltage (value of one small scale div) should be as small as possible (at least 0.2 volts).

As most of the cells have voltages less than 2 volts, the meters used to measure cell voltage should be free from errors. One of the most common errors in meters is the **Mechanical Zero error**. This error is caused due to the mechanical movements involved in the meters. This error in meters is correctable. The steps involved to correct this error is called *Mechanical zero setting of meters*.

All meters will have a screw on it as shown in Fig 6. Keeping the terminals of the meter open, the screw is turned slowly to bring the pointer exactly to 0 position on the meter scale. This means, with no voltage applied, the meter is made to show exactly zero volts.



Care has to be taken while turning this screw as this screw is directly connected with the sensitive and delicate meter movement. Turning the screw in large amounts or in jerks may damage the meter movement permanently making the meter unusable.

Before using a meter for measurements, it is necessary to check if the meter needle is moving freely over the graduated scale. There are possibilities that the meter movement may be sticky due to dust collection on the meter movement or due to the bent pointer needle.

A simple way to check sticky pointer/meter movement is to hold the meter in hand and tilt the meter back and forth gently, checking for the free movement of the pointer. If the pointer is not moving freely, it is advised not to use that meter for making measurements.

Voltmeters used for measuring DC voltages will have their input terminals marks +ve and -ve. For making voltage measurement, the +ve terminal of the meter must be connected to the +ve terminal of the battery and the -ve terminal of the meter to the -ve terminal of battery. If the terminals are reversed, the meter deflects below zero. This may cause temporary or sometimes permanent damage to the meter movement.

The following precautions should always be taken while making voltage measurement of cells or other voltage sources;

- 1) Check if the right type and range of meter is being used for measurement.

Example: To measure cell/battery voltages use DC or AC/DC voltmeters only.

- 2) Check if the maximum and minimum measurable voltage of the meter is as per requirement.

Example: To measure cell voltages of around 1.5volts, use a voltmeter having a range of 2 volts but not more than 5 volts. The value of each small scale division of the meter should be atleast 0.2 volt or less.

- 3) Make it a practice to connect red coloured meter prod to the +ve terminal of the meter and black coloured prod to the -ve terminal. This avoids confusion and prevents damage to meter.

- 4) Never measure the current of a cell by connecting an ammeter directly across its +ve and -ve terminals. This will not only damage the meter but will discharge the cell heavily and make the cell weak. Use a 1 KW resistor in series for measuring current driving capacity of a cell.

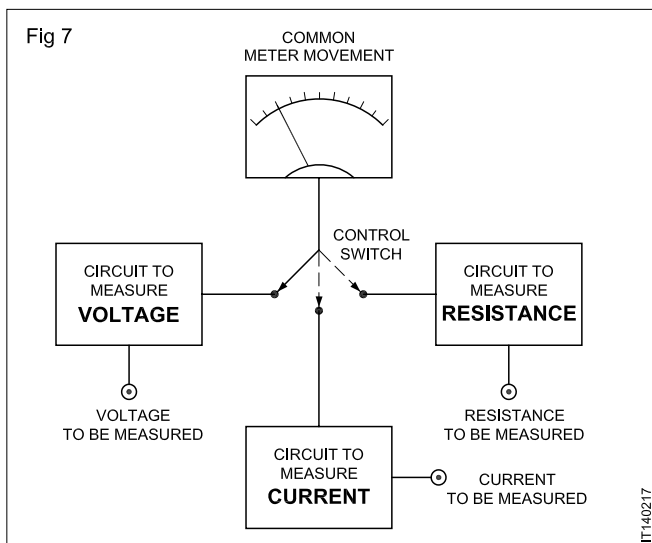
MULTI-METER

The voltmeter, ammeter and ohmmeter although used for measuring different electrical parameters - all these meters use the same basic type of meter movement called the D'Arsonval movement (named after its inventor). The difference between these instruments while measuring different electrical parameters are

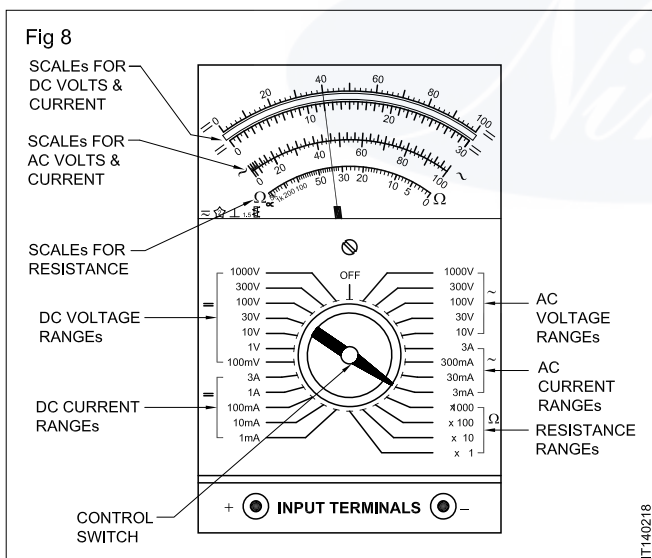
- the circuit associated with the meter movement
- the graduations/markings on its meter scale.

NOTE: Meter movement is an electro mechanical arrangement used to deflect the meter pointer. Details of types of meter movement is discussed in further lessons.

It is, therefore, possible to design an instrument having a common meter movement, with three different circuits, for measuring voltage, current, resistance. A control switch will be necessary to connect any one of the three circuits to the common meter movement as shown in Fig 7.



Such an instrument which can be used to measure voltage, current and resistance is called a **Volt - Ohm - milliammeter** or **VOM** or more commonly known as **multi-meter**. Fig 8 shows a typical multi-meter.



NOTE: The multimeter used in your institute may have a different front panel design than shown in Fig 8. But the function will remain almost same.

In the multimeter shown in Fig 8, following steps are to be followed to measure dc voltages.

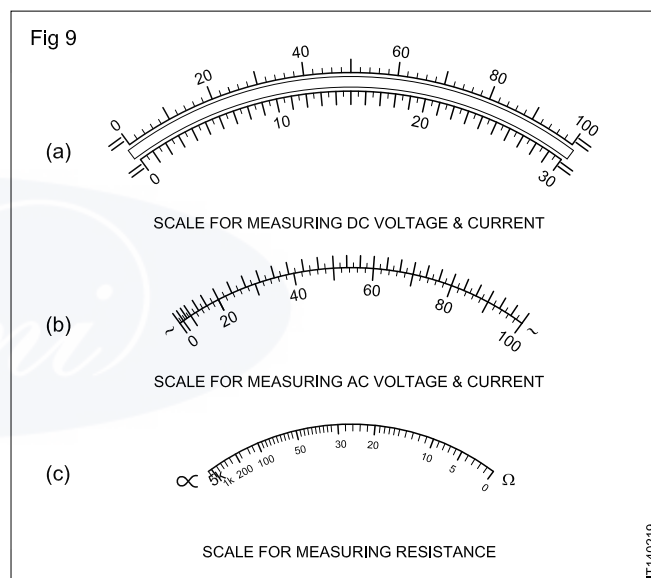
- First : Know the approximate voltage to be measured.
- Second: Put the control switch to the most suitable dc volts range.

While measuring unknown voltage levels, always set the range switch to the highest available range and work down from there. This avoids damage to meter.

- Third: Identify the most suitable scale on which the readings are to be taken corresponding to the chosen range.
- Fourth: Find the value of one small scale division of the scale.
- Fifth: Give the voltage to be measured to the meter using the meter probes with correct polarity.
- Sixth: Read the indicated voltage on the scale.

Identifying the scale to take readings

In almost all meters, the symbol of the electrical parameter for which the particular scale is to be used will be given either at the beginning, or at the end, or at both places as shown in Figs 9a, b, c.



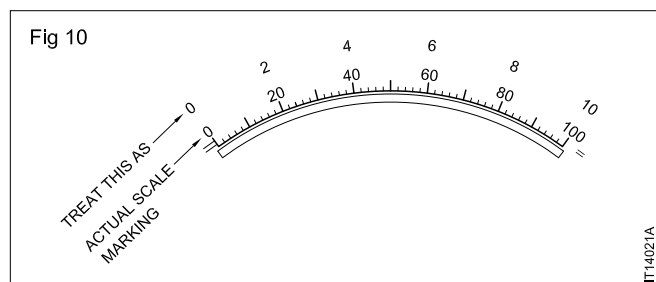
The symbol “-” at the beginning and end of scale shown in Fig 9a indicates that this scale should be used for dc voltages and current.

Similarly, Fig 9b shows the scale to be used for ac voltage and current and Fig 9c for resistance.

In the scale meant for taking readings for a particular parameter(say dc voltage) may itself have more than one scale for the convenience of taking readings in different range positions.

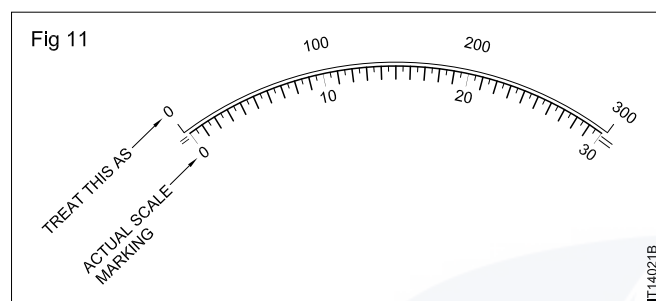
For example, the scale to be used for dc voltages shown in Fig 9a itself has two scales 0-100 and 0-30. Of these two scales, depending upon the range switch position selected, readings will have to be taken on any one scale.

Example 1: In the multimeter at Fig 8, if the range switch is put to 10 V position, choose 0-100 scale. Treat the maximum marking 100 as 10 as shown in Fig 10.



Here each SSD on meter scale corresponds to 0.2 volts and hence it is accurate as well as easy to take the readings.

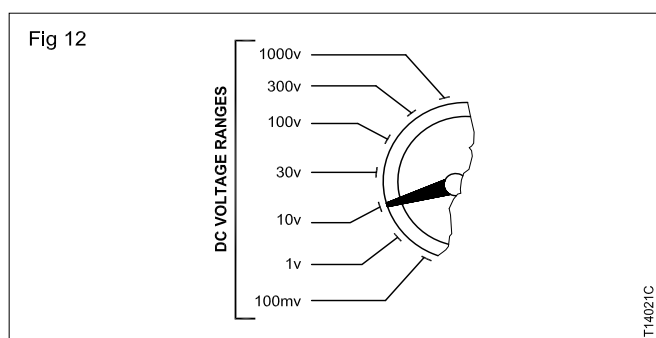
Example 2: Referring to Fig 8, if the range switch is put to 300 volts position, choose 0-30 scale. Treat the maximum marking 30 as 300 as shown in Fig 11.



Using multi-meters as voltmeter, ammeter and ohmmeter

Multimeter as DC VOLTMETER

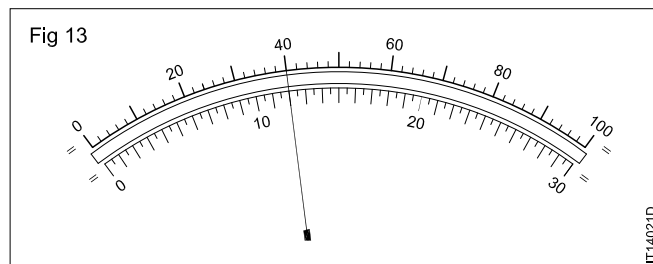
To measure a wide range of voltages the multi-meter will have several ranges under **dc volts** as shown in Fig 12. This is similar to a multi-range voltmeter. Depending on the quantity of voltage being measured the control switch can be put to the most suitable range position. Corresponding to the range selected the most suitable scale has to be chosen for taking the readings.



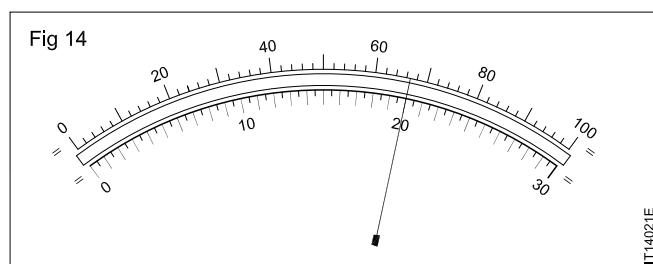
Example 1: Referring to Fig 12, to measure the voltage of a 1.5 V dry cell, the control switch should be put to 10 V position. This is because, the battery voltage is greater than 1 V and less than 10 V.

Since the range selected is 10 V, the most suitable scale to take reading is the scale with the marking 0 - 100.

Hence, with the control switch in 10V position, if the pointer indication is as in Fig 13 the voltage indicated will be 4 volts.



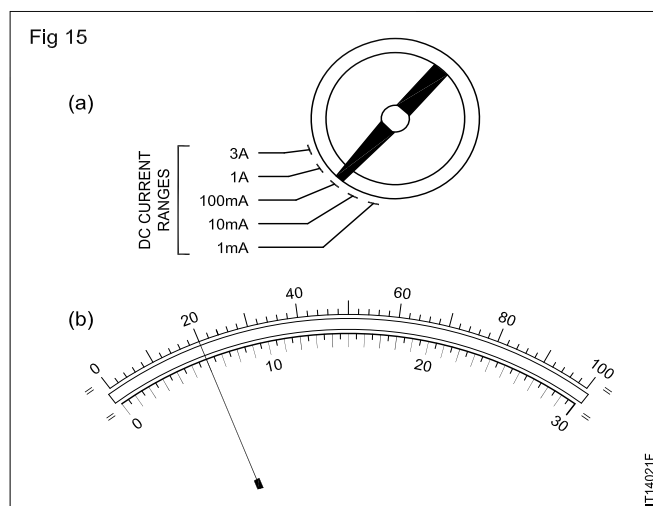
Example 2: With the control switch in 300 volts position, if the pointer indication is as in Fig 14, the voltage indicated will be 200 volts.



Multi-meter as DC AMMETER

To use a multi-meter as a dc ammeter, the control switch should be put to the most suitable dc current position as shown in Fig 15a. This depends upon the quantity of dc current to be measured. *If in doubt put to maximum range position.*

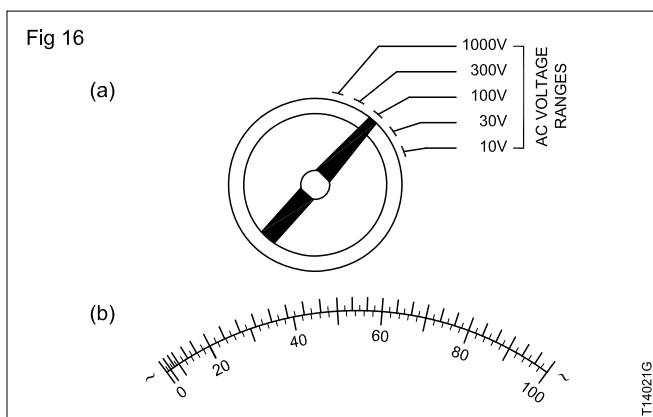
Choose the scale shown in Fig 15b to read the dc current.



Multi-meter as AC VOLTMETER

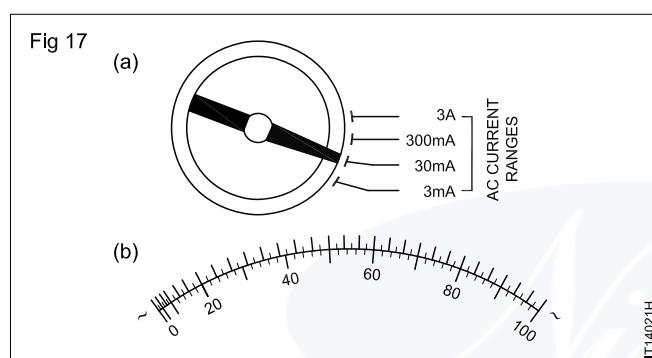
To use a multi-meter as an ac voltmeter, the control switch should be put to ac volts position as shown in Fig 16a.

Choose the scale shown in Fig 16b to read the ac voltages.



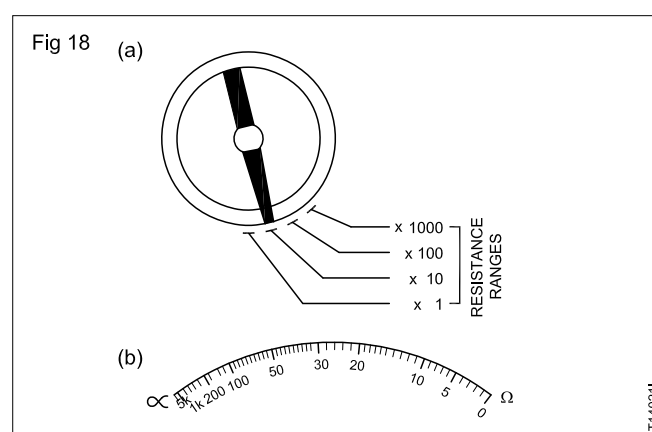
Multi-meter as AC AMMETER

Put the control switch to ac current position and the most suitable current range as shown in Fig 17a and choose scale shown in Fig 17b.

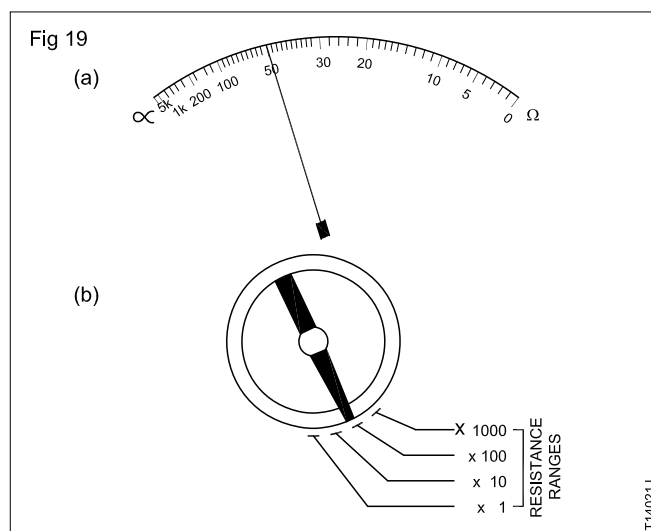


Multi-meter as OHM METER

Put the control switch to the most suitable OHMS (W) position as shown in Fig 18a and use the scale shown in Fig 18b.



The range switch in resistance range also has different range positions to enable the measurement of a wide range of resistance values. But unlike in voltage and current range positions, these ranges are marked **x1000**, **x100** and so on. This means, whatever is the value of resistance pointed by the meter needle should be multiplied by this factor.



Example: If the meter needle shows a reading as shown in Fig 19a with the control switch in position as in Fig 19b the value indicated by the meter will be;

Meter reading = $(50) \times 100 = 5000 \text{ W}$ or 5 KW .

Resistance zero setting

One important thing to note in the scale for measuring resistance is that, the zero of the scale is on the right hand end of scale. When the meter prods are shorted, the resistance across the meter terminals is zero and the meter should indicate 0W. Sometimes the pointer may not be exactly at zero. Then, it is required to set the pointer to indicate zero. To set the pointer to zero short the meter prods and turn the *Resistance zero set knob* provided on the meter such that the pointer goes to zero. This procedure is called *resistance zero setting* of the meter.

Note that Resistance zero setting is different from Mechanical zero setting carried out in this exercise.

Mechanical zero setting is to adjust errors arising due to the mechanical parts of the meter whereas **Resistance zero setting** is to adjust the meter full scale deflection due to low voltage of battery inside the meter. This adjustment is required to be carried out only when using the meter to measure resistance.

While using a multi-meter as ohmmeter, whenever you change the multiplier range position, it is advised to carry out the **Resistance zero setting** of the meter to get accurate readings.

There are several manufacturers of multi-meters and the meters differ from one another in the provided ranges and meter scales. Therefore, before using any multi-meter it is worth spending some time in studying the various ranges and scales available on the meter.

In some low cost multi-meters there will be no control knob for changing the meter range. Instead they will have a common -ve socket where the black meter prod should always be plugged in. Depending upon the range required, the +ve meter prod should be plugged into the socket provided for each range.

TYPES OF CELLS AND THEIR APPLICATIONS

Usual Name/Types	Classification	Negative Electrode (Anode)	Positive Electrode (Cathode)	Electrolyte	Rated output voltage	Available sizes	Applications
Carbon-Zinc (usually called Leclanche cell)	Primary	Zinc	MnO ₂ /C	Mixture of NH ₄ Cl and ZnCl ₂	1.5V	D, C, B, A, AA, AAA	Flash lights, radio, tape batteries and for general purpose.
Carbon-Zinc (Zinc chloride cell)	Primary	Zinc	MnO ₂ /C	Zinc chloride	1.5V	D, C, B, A, AA, AAA	Electric shaver, electric knives transmitters, cordless drills, tools etc.
Alkaline-Manganese dioxide cell	Primary & rechargeable	Zinc	Manganese dioxide	Aqueous solution of potassium hydroxide	1.5V	D, C, AA	Camera cranking, radio controlled toys, radios & tape recorders.
Mercuric oxide cell	Primary	Zinc	Mercuric oxide	Aqueous solution of potassium hydroxide or sodium hydroxide	1.35V	C, B, AA and button cells	Cameras, watches, hearing aids, calculators etc.
Silver oxide cell	Primary	Zinc	Ag ₂ O	Aqueous solution of potassium hydroxide or sodium hydroxide	1.5V	Button cells	Hearing aids, digital wrist watch, micro-lamps, lights, meters etc.
Nickle-Cadmium	Rechargeable	Cadmium	Nickel hydroxide	Aqueous solution of potassium hydroxide	1.2V	All sizes of cyclin- drical, rectangular and button cells	Portable equipments like radio, tapes etc., rechargeable flash lights, emergency light etc.
Lithium Manganese	Primary	Lithium	Iodine/metallic oxides, sulphides	Organic, inorganic water	3V to 6 V	Medium to large button cells	Electronic watches, cal- culators, heart pacemaker life support equipments & communication equipments.

Series & parallel connection of batteries

Objectives: At the end of this lesson you shall be able to

- state and describe the principle and construction of a primary cell
- list the common primary cells in use
- explain the effect of the internal resistance of a cell.

Cell: A cell is an electrochemical device consisting of two electrodes made of different materials and an electrolyte. The chemical reaction between the electrodes and the electrolyte produces a voltage.

Cells are classified as

- dry cells
- wet cells.

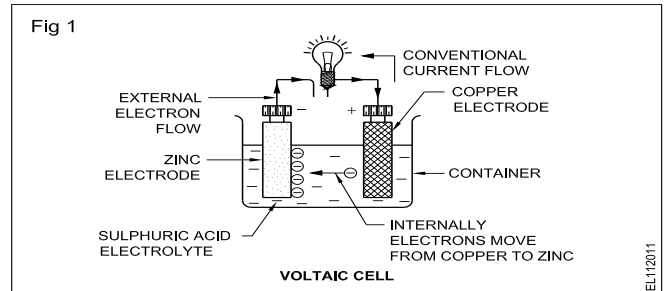
Cells and batteries are classified as either dry or wet. Historically, a dry cell is one that has a paste or gel electrolyte. It is semi-sealed and could be used in any position. With newer designs and manufacturing techniques, it is possible to completely (hermetically) seal a cell. With complete seals and chemical control of gas build-up, it is possible to use liquid electrolytes in dry cells. Today the term 'dry cell' refers to a cell that can be operated in any position without electrolyte leakage.

Wet cells are cells that must be operated in an upright position. These cells have vents to allow the gases generated during charge or discharge to escape. The most common wet cell is the lead-acid cell.

Primary cells: Primary cells are those cells that are not rechargeable. That is, the chemical reaction that occurs during discharge is not easily reversed. When the chemicals used in the reactions are all converted, the cell is fully discharged. It must then be replaced by a new cell. Included in the primary cell category are the following types.

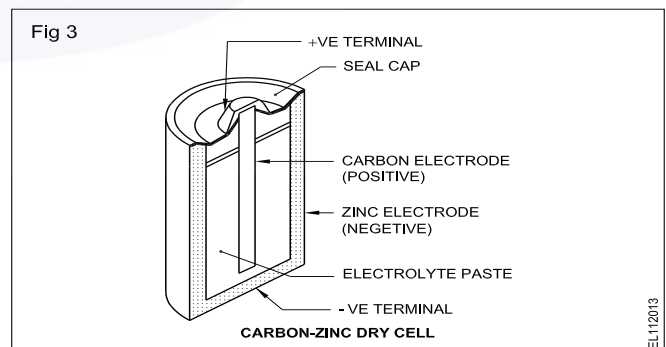
- Voltaic cell
- Carbon-zinc (Leclanche cell) (Dry cell)
- Alkaline
- Mercury
- Silver oxide
- Lithium

Simple voltaic cell: A voltaic cell uses copper and zinc as the two electrodes and sulphuric acid as the electrolyte. When placed together a chemical reaction occurs between the electrodes and the sulphuric acid. This reaction produces a negative charge on the zinc (surplus of electrons) and a positive charge on the copper (deficiency of electrons). If an external circuit is connected across the two electrodes, electrons will flow from the negative zinc electrode to the positive copper electrode. (Fig 1) The electric current will flow as long as the chemical action continues. In this type of cell the zinc electrode is eventually consumed as part of the chemical reaction.



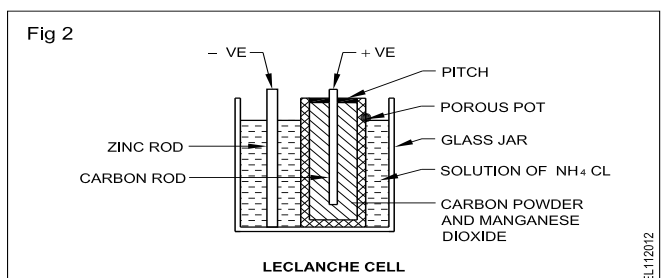
The voltaic cell is also known as a wet cell because it uses a liquid solution for the electrolyte.

Leclanche cell (Carbon-zinc cells): The container of this cell is a glass jar. The jar contains a strong solution of ammonium chloride (NH_4Cl). This solution is an alkali and acts as the electrolyte. A porous pot is placed at the centre of the glass jar. This porous pot has in it a carbon rod surrounded by a mixture of manganese dioxide (MnO_2) and powdered carbon. The carbon rod forms the positive electrode of the cell and MnO_2 acts as the de-polarizer. A zinc rod is dipped in the solution in the jar and acts as the negative electrode. (Fig 2)



Dry cell (Carbon-Zinc cell): The danger of spilling the liquid electrolyte from a Leclanche type of cell led to the invention of another class of cells called dry cells.

The most common and least expensive type of a dry cell battery is the carbon-zinc type. (Fig 3) This cell consists of a zinc container which acts as the negative electrode. In the centre is a carbon rod which is the positive electrode.

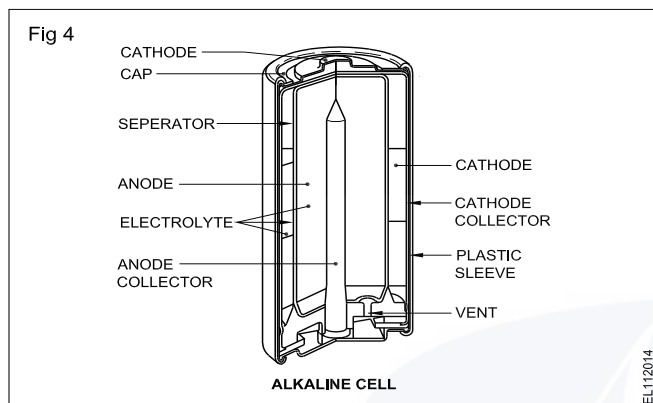


The electrolyte takes the form of a moist paste made up of a solution containing ammonium chloride. As with all primary cells, one of the electrodes becomes decomposed as part of the chemical reaction. In this cell the negative zinc container electrode is the one that is used up. As a result, cells left in equipment for long periods of time can rupture, spilling the electrolyte and causing damage to the neighbouring parts.

Carbon-zinc cells are produced in a range of common standard sizes. These include 1.5 V AA, C, D cells.

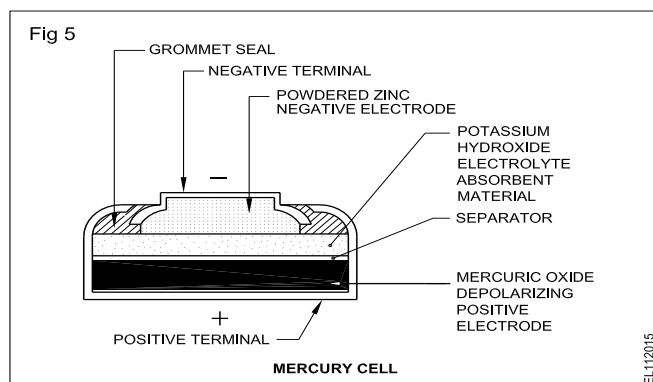
(AA Pen type cell, 'C' medium size, 'D' large/economy size.)

Alkaline cells: Alkaline cells use a zinc container for the negative electrode and a cylinder of manganese di-oxide for the positive electrode. (Fig 4) The electrolyte is made up of a solution of potassium hydroxide or an alkaline solution.



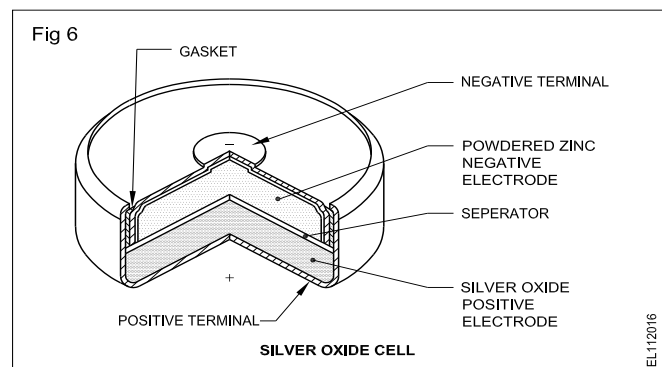
Alkaline cells are produced in the same standard sizes as carbon-zinc cells but are more expensive. They have the advantage of being able to supply large currents for a longer period of time. For example, a standard 'D' type 1.5 V alkaline cell has a capacity of about 3.5 A.h compared with about 2 A.h for the carbon-zinc type. A second advantage is that the alkaline cell has a shelf life of about two and a half years as compared to about 6 to 12 months for the carbon-zinc type.

Mercury cells: Mercury cells are most often used in digital watches, calculators, hearing aids and other miniature electronic equipment. They are usually smaller and are shaped differently from the carbon-zinc type. (Fig 5)



The electrolyte used in this cell is alkaline and the electrodes are of mercuric oxide (cathode) and zinc (anode).

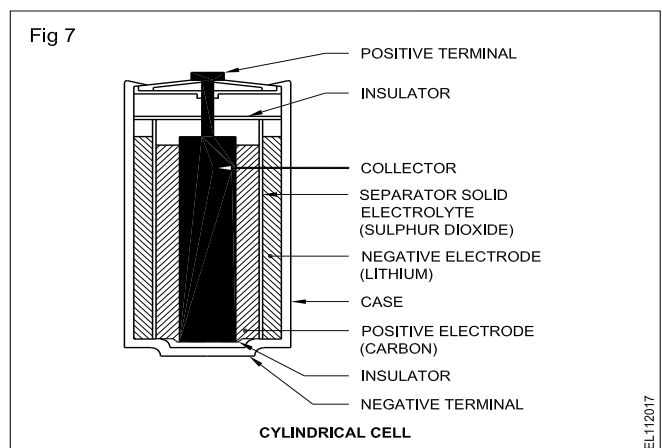
Silver oxide cells: Silver oxide cells are much like mercury cells. However, they provide a higher voltage (1.5 V) and they are made for light loads. The loads can be continuous, such as those encountered in hearing aids and electronic watches. Like the mercury cell, the silver oxide cell has good energy-to-weight and energy-to-volume ratios, poor low-temperature response, and flat output voltage characteristics. The structures of the mercuric and silver oxide cells are very similar. The main difference is that the positive electrode of the silver cell is silver oxide instead of mercuric oxide. Fig 6 shows the cross-section of a silver oxide cell.



Lithium cells: The lithium cell is another type of primary cell. (Fig 7) It is available in a variety of sizes and configurations. Depending on the chemicals used with lithium, the cell voltage is between 2.5 and 3.6 V. Note that this voltage is considerably higher than in other primary cells. Two of (Fig 7) the advantages of lithium cells over other primary cells are:

- longer shelf life - up to 10 years
- higher energy-to-weight ratios up to 350 Wh/Kg.

Lithium cells operate at temperatures ranging from -50 to +75°C. They have a very constant output voltage during discharge.



Uses: Primary cells are used in electronic products ranging from watches, smoke alarms, cardiac pacemakers, torches, hearing aids, transistor radios etc.

Internal resistance: The output voltage from a cell varies as the load on the cell changes. Load on a cell refers to the amount of current drawn from the cell. As the load increases, the voltage output drops, and vice versa. The change in output voltage is caused by the internal resistance of the cell.

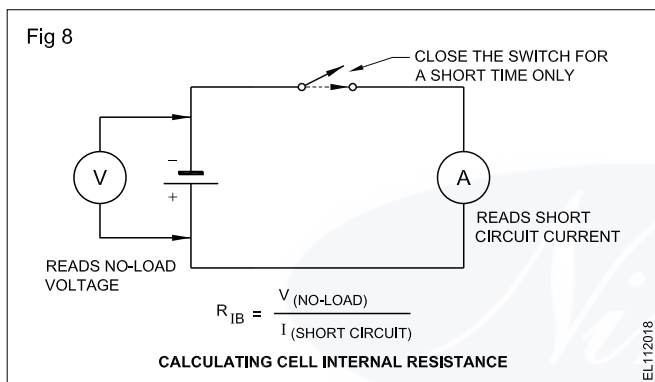
Since materials from which the cell is made are not perfect conductors, they have resistance. Current flowing through the external circuit also flows through the internal resistance of the cell.

According to Ohm's law, a current flowing through a resistance (either external or internal) results in a voltage drop ($V = IR$). Any voltage developed across the internal resistance is not available at the terminals of the cell.

The voltage at the terminals is the voltage due to the chemical reactions minus the voltage dropped across the internal resistance. The terminal voltage of a cell, therefore, depends on both the internal resistance of the cell and the amount of load current.

In some applications, the changes in cell terminal voltage are so small that they make no practical difference. In other applications, the changes are very noticeable. For example, when an automobile engine is started, the battery output voltage changes from about 12.6 to 8V.

The method of calculating the internal resistance of a cell is shown in Fig 8.



As a cell discharges, its internal resistance increases. Therefore its output voltage decreases for a given value of load current.

Defects of a simple cell: With a simple voltaic cell, the strength of current gradually diminishes after some time. This defect is mainly due to two causes.

- Local action
- Polarisation

Local action: In a simple voltaic cell, bubbles of hydrogen are seen to evolve from the zinc plate even on open circuit. This effect is termed local action. This is due to the presence of impurities like carbon, iron, lead, etc. in the commercial zinc. This forms small local cells on the zinc plate and reduces the strength of current of the cell.

The local action is prevented by amalgamating the zinc plate with mercury. To do so, the zinc plate is immersed in dilute sulphuric acid for a short time, and afterwards, mercury is rubbed over its surface.

Polarisation: As current flows, bubbles of H_2 evolve at the copper plate on which they gradually form a thin layer. Due to this the current strength falls and finally stops altogether. This effect is called the polarization of the cell.

Polarisation can be prevented by using some chemicals which will oxidize the hydrogen to water before it can accumulate on the plate. The chemicals used to remove polarisation are called de-polarisers.

COMPARISON OF COMMON PRIMARY CELLS

	Carbon -Zinc	Alkaline-Manganese	Mercury	Silver oxide
Negative, anode Positive, cathode Electrolyte	Zinc Carbon Ammonium Chloride	Zinc Manganese dioxide Potassium hydroxide	Zinc Mercuric oxide Potassium hydroxide	Zinc Silver oxide Potassium hydroxide
Nominal voltage, volts Max. rated current amperes	1.5 2-30	1.5 0.05-20	1.35 or 1.4 0.003-3	1.5 0.1
Energy output Watt-hrs Ampere-hours	22 2.0	35 3.5	46 6.0	50 8.0
Temperature range Storage °F Operating °F	-40 to 120 20 to 130	-40 to 120 -5 to 160	-40 to 140 -5 to 160	-40 to 140 -5 to 160
Shelf life in months at 68°F to 80% initial capacity Shape of discharge curve	6 to 12 Sloping	30 to 36 Sloping	30 to 36 Flat	30 to 36 Flat

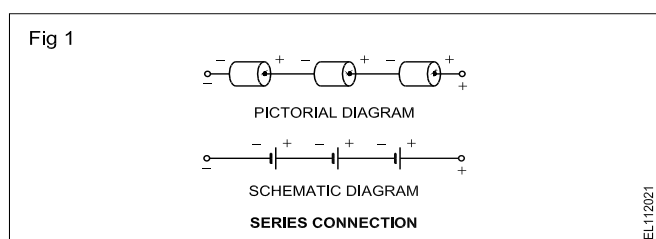
Grouping of cells

Objectives: At the end of this lesson you shall be able to

- calculate the total voltage for the given cells connected in series
- calculate the cells required in series for a given voltage
- state the precautions to be observed while connecting cells in series
- calculate the cells in parallel required to get the rated ampere-hour output
- state the precaution to be observed while connecting cells in parallel
- calculate the cell groups required in series parallel groups to get required voltage and ampere-hour rating
- state the method of testing cells.

Series and parallel cell connections: Often an electric circuit requires a voltage or current that a single cell is not capable of supplying alone. In this case it is necessary to connect groups of cells in various series and parallel arrangements.

Series cell connections: Cells are connected in series by connecting the positive terminal of one cell to the negative terminal of the next cell. (Fig 1)

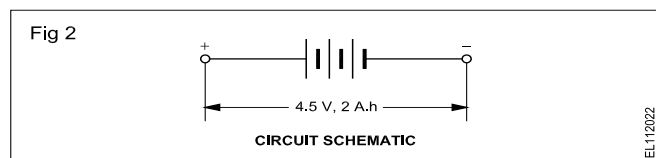


Identical cells are connected in series to obtain a higher voltage than is available from a single cell. With this connection of cells, the output voltage is equal to the sum of the voltages of all the cells. However, the ampere hour (A.h) rating remains equal to that of a single cell.

Example: Suppose three 'D' flashlight cells are connected in series. (Fig 2) Each cell has a rating of 1.5 V and 2 A.h. The voltage and ampere hour rating of this battery would be:

$$\begin{aligned} \text{V Battery} &= \text{V per cell} \times \text{No. of cells} \\ &= (1.5\text{V}) (3) \\ &= 4.5 \text{ V} \end{aligned}$$

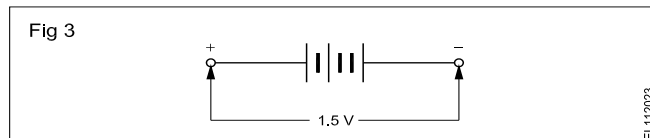
$$\begin{aligned} \text{A.h Battery rating} &= \text{A.h rating of 1 cell} \\ &= 2 \text{ A.h} \end{aligned}$$



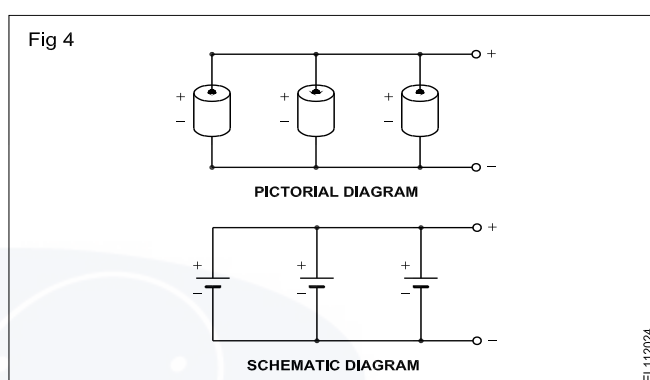
If, by mistake, one cell connection is reversed in a series group, its voltage will oppose that of the other cells. This will produce a lower than expected battery output voltage.

Example: Suppose that one of the three 'D' flashlight cells of the previous example is connected in reverse, (Fig 3) the output voltage then would be:

$$\begin{aligned} \text{V Battery} &= (1.5\text{V}) + (1.5\text{V}) - (1.5\text{V}) \\ &= (3\text{V}) - (1.5\text{V}) = 1.5\text{V}. \end{aligned}$$



Parallel connection: Cells are connected in parallel by connecting all the positive terminals together and all the negative terminals together as in Fig 4.

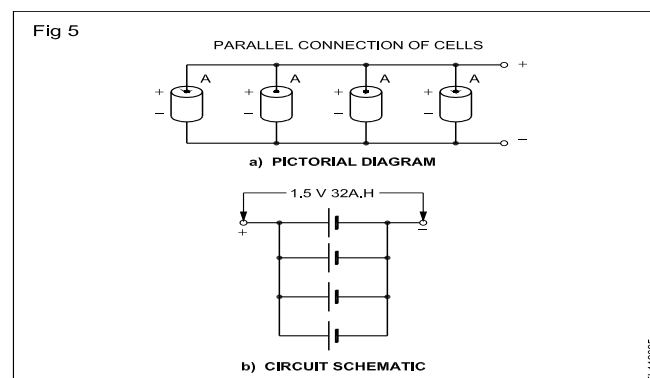


Identical cells are connected in parallel to obtain a higher output current or ampere-hour rating. With this connection of cells, the output ampere hour rating is equal to the sum of the ampere hour ratings of all the cells. However, the output voltage remains the same as the voltage of a single cell.

Example: Suppose four cells are connected in parallel. (Fig 5) Each cell has a rating of 1.5 V and 8 A.h. The voltage and ampere-hour rating of this battery would be:

$$\begin{aligned} \text{V Battery} &= \text{V rating of 1 cell} \\ &= 1.5 \text{ V} \end{aligned}$$

$$\begin{aligned} \text{A h Battery rating} &= \text{A h rating per cell} \times \text{no. of cells} \\ &= (8 \text{ A h}) (4) \end{aligned}$$



If, by mistake, a cell connection is reversed in a parallel group, it will act as a short circuit. All cells will discharge their energy through this short circuit path. Maximum current will flow through the short circuit and the cells may be permanently damaged.

Series parallel connection: Sometimes the requirements of a piece of equipment exceed both voltage and ampere hour rating of a single cell. In this case a series-parallel grouping of cells must be used.

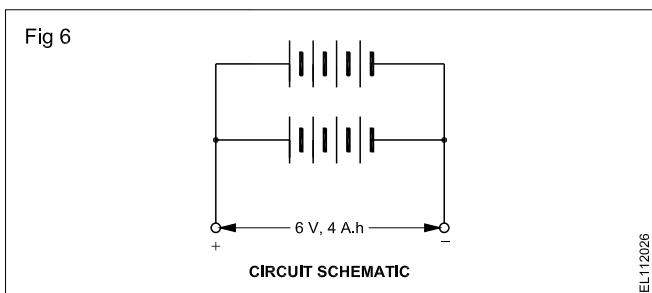
The number of cells that must be connected in series to have voltage rating is calculated first and then the number of parallel rows of series connected cells is calculated for required ampere-hour rating.

Example: Suppose a battery operated circuit requires 6 V and a capacity of 4 A h. (Fig 6) Cells rated at 1.5 V and 2 A h are available to do the job. The required arrangement of cells would then be:

$$\begin{aligned}\text{No. of cells in series} &= \left(\frac{V_{\text{required}}}{V_{\text{per cell}}} \right) \\ &= \frac{6 \text{ V}}{1.5 \text{ V}} = 4 \text{ cells}\end{aligned}$$

$$\begin{aligned}\text{No. of parallel rows} &= \left(\frac{A.h_{\text{required}}}{A.h_{\text{per cell}}} \right) \\ &= \frac{(4 \text{ A.h})}{2 \text{ A.h}} = 2 \text{ rows.}\end{aligned}$$

(Refer Fig 6)



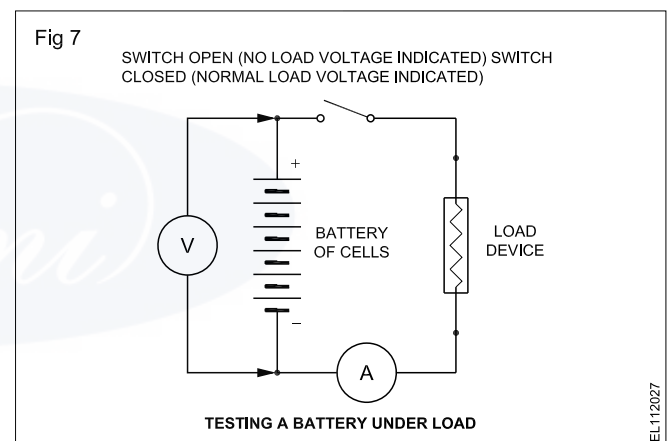
When connecting groups of cells or batteries in parallel, each group must be at the same voltage level. Paralleling two batteries of unequal voltage levels sets up a difference of potential energy between the two. As a result, the higher voltage battery will discharge its current into the other battery until both are at equal voltage value.

Testing primary cells or batteries: A visual inspection will tell you little about the useful life of a cell or battery unless it has deteriorated to the point where acid is spilling from the case.

A no-load voltage test of the cell or battery is another indication of cell or battery life. This test requires the cell or battery to deliver only a very small amount of current required to operate the voltmeter.

The best method that is used to check a cell or battery is an in-circuit test of the cell or battery voltage with the normal load connected to it. (Fig 7) A substantial drop in cell or battery voltage, when normal load is applied, indicates a bad cell or battery.

$$= 32 \text{ A h.}$$



Charging techniques of secondary batteries

Objectives: At the end of this lesson you shall be able to

- state the important differences between primary and secondary cells
- state the chemical reaction during charging and discharging of a lead-acid wet cell.
- state the meaning of the capacity of and effect of temperature of batteries
- state the importance of specific gravity of the electrolyte used in wet type lead-acid batteries
- state the use of high rate discharge tester and the meaning of the terms topping up
- list care and routine maintenance of lead-acid batteries
- state type of methods and the difference between constant current and constant voltage battery charging
- state the terms, trickle charging and float charging
- list the precautions to be taken while charging lead-acid batteries.

INTRODUCTION

Recall the following points.

- 1 Batteries are sources of DC electrical energy.
- 2 A battery is an electro chemical system that converts chemical energy into electrical energy.
- 3 A battery consists of combinations of small units called cells.
- 4 Cells can be broadly divided into two main categories, namely, primary cells and secondary cells.
- 5 Depending upon the type of electrolyte used, the batteries can also be classified as dry cells or wet cells.
- 6 Different types of primary cells are available.

Example: Zinc-carbon, zinc chloride, manganese dioxide (Alkaline), mercuric oxide, silver oxide, lithium and so on.

Once the charge stored in the primary cells is used up, then they have to be thrown off. In other words they cannot be recharged for reuse. This makes the use of primary cells expensive. Also primary cells cannot be used in high load current requirements. In other words the capacity of a battery, which is an indication of the total usable energy available from the battery, stated in ampere-hour (Ah) is very low in case of primary cells/batteries.

In applications where high values of load current are necessary, *secondary batteries* are preferred over primary batteries.

Secondary batteries

Secondary batteries are made of small units known as cells. The main difference between a primary and a secondary cell is that a secondary cell can be recharged. This is because the type of chemicals used in a secondary cell is such, the chemical reaction is reversible.

When a secondary cell is supplying current to a load, the cell is said to be *discharging*. This discharging current gradually neutralizes the separated positive and negative charges at the electrodes (Anode and Cathode).

On the other hand, when current is supplied to a cell, the charges get re-formed on the electrodes due to reverse chemical reaction. This action is known as *charging* the cell. For charging a cell, the charging current is supplied by an external dc voltage source, with the cell behaving as a load.

The process of discharging and recharging is called *cycling* of the cell. As long as the cell is in good condition the discharge and charge cycles can be repeated several hundred times.

Since a secondary cell can be recharged, in other words the charges restored, these cells are called *storage cells*.

The most common type of secondary cell is the *Lead-acid cell*. A battery consisting of a combination of such cells is called *Lead-acid battery*. Lead-acid batteries are commonly used in automobiles such as cars, buses and lorries etc.,

Lead-acid, wet type cells

Lead-acid secondary batteries made of lead-acid are used in almost every automobile, for starting the engine. These batteries supply load current of 100 to 400A to the starter motor of automobiles.

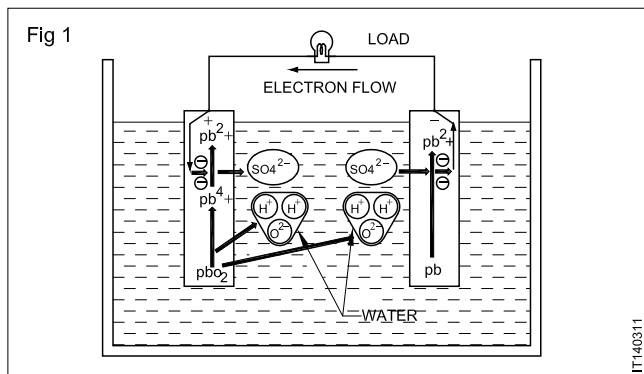
The nominal voltage of a lead-acid cell is 2.2 V. By connecting three or six cells in series, batteries of 6V or 12V is obtained.

Principle of chemical action

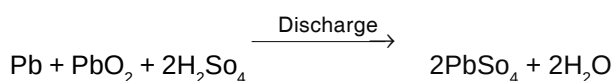
A fully charged lead-acid cell has a lead peroxide (PbO_2) positive electrode, which will be reddish brown in colour and a gray spongy lead (Pb) as the negative electrode. These two electrodes are immersed in an electrolyte which is a diluted solution of sulphuric acid (27% sulphuric acid) having a specific gravity of 1.3. Such a cell produces an output of 2.2 V.

Discharging of lead-acid cells

The chemical action that take place during the discharging of a lead-acid cell is shown in Fig 1.



During discharge, the lead (Pb) in both the electrodes react with sulphuric acid (H_2SO_4) to displace hydrogen and form lead sulphate (PbSO_4). This lead sulphate, a whitish material, is somewhat insoluble and hence gets partially coated on both positive and negative plates. Since both plates approach the same material (PbSO_4) chemically, the potential difference between these plates begins to decrease. At the same time, the combining of oxygen in the lead peroxide (PbO_2) with the hydrogen atoms of the electrolyte forms water (H_2O) as shown in the equation given below,

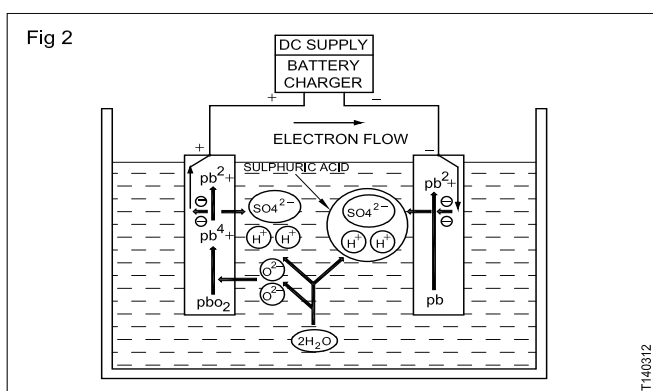


It can be seen from the discharging equation that as the battery discharges (delivers energy to a load), the sulphuric acid solution becomes weaker (more and more diluted) with its specific gravity approaching 1.0.

The coating of whitish lead sulphate on the electrodes and the decrease in specific gravity of the electrolyte makes the voltage of the cell to drop off. Also, the internal resistance of the cell rises due to the sulphate coating on the plates.

Charging of lead-acid cells

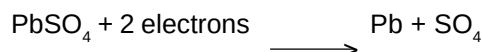
The chemical reaction that take place during charging of a lead-acid cell is shown in Fig 2.



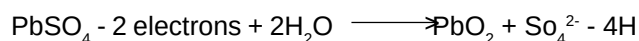
When a battery charger, having an output voltage (2.5V) which is slightly higher than the nominal voltage of the cell (2.2V), is connected as shown in Fig 2, the direction of ionic flow gets reversed (refer to Fig 1 for the discharging direction).

The electrical energy supplied by the charger causes the recombination of lead sulphate (PbSO_4) with hydrogen ions in the electrolyte. Therefore, the excess water is removed from the electrolyte solution. As the electrolyte returns to its normal strength of sulphuric acid (27%) and the plates return to their original form of lead peroxide and spongy lead, the voltage across the electrodes returns to its nominal value of 2.2V. The chemical action involved during charging can be represented by the following equation;

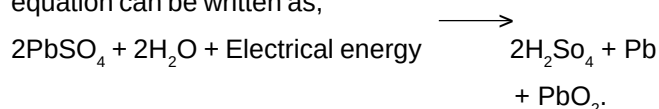
At the negative pole:



At the positive pole:



As the above reactions take place simultaneously, the equation can be written as,



IMPORTANT TIPS: It should be noted that, to charge a lead-acid battery of 12 V (2.2 V x 6 cells), the output voltage of the battery charger used for charging should be between 14.1 V to 15 V, and, its current rating not larger than 30 A. Charging batteries at excessively higher currents can cause **boiling** of the electrolyte. This reduces the liquid level in the battery and causes **backing** and **crumbling** of the electrodes, thus reducing the life of the cells and hence the battery.

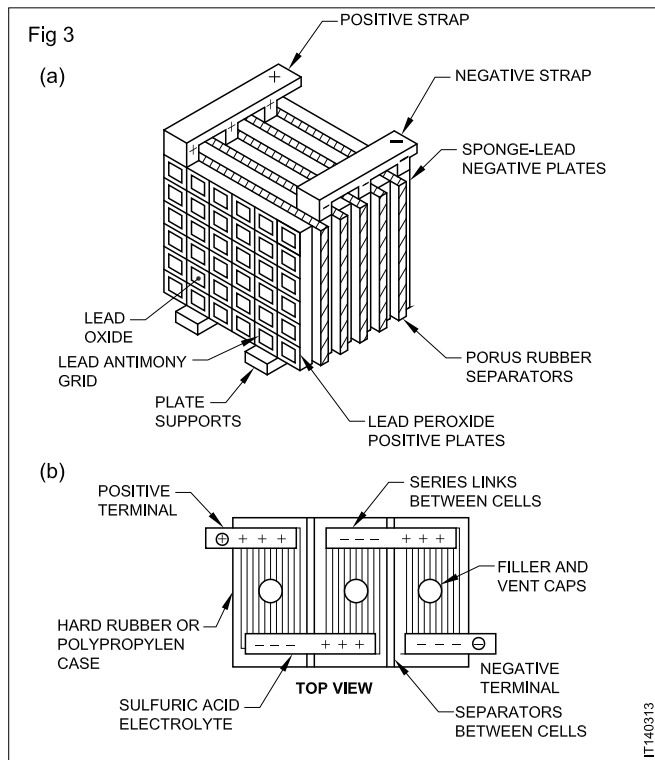
The lead sulphate (PbSO_4) which gets coated on the +ve and -ve plates tends to harden into an insoluble salt over a period of time. Hence, it is recommended to fully recharge a battery even if it is not used for quite some time.

Construction of lead-acid batteries

Fig 3 shows the principle behind the construction of commercial lead acid batteries.

Although in Figs 1 and 2, the lead-acid cell electrodes were shown as single plates, in a practical cell, it will not be the case. To increase the surface area and current capacity, a number of positive and negative plates are interleaved and separated by porous rubber sheets as shown in Fig 3a. All the positive plates are electrically connected, and all the negative plates are electrically connected. These parallel connections yield a higher current capacity of the cell with an overall cell output voltage of 2.2V. Several such cells can be connected in series to obtain the required battery voltage. For example, Fig 3b shows three such cells connected in series to produce a 6 volts Lead acid battery.

In lead-acid batteries, since hydrogen gas is produced during recharging, vents (holes) are provided on the battery compartment to let hydrogen and water vapour escape into free air. The vents also help in adding distilled water to the cells to compensate the water evaporated from the electrolyte.



For further details on the construction and manufacturing techniques of lead acid batteries refer reference books listed at the end of this unit.

Current rating of Lead acid batteries

The current rating of a lead acid battery is usually given in ampere-hour (Ah) units, based on an 8 hour discharge period. In other words, batteries are rated in terms of how much discharge current they can supply for a specified period of time (often 8 hours). During this time, the cell's output voltage must not drop below 1.7 volts. Typical Ah values of automobile batteries range from 60 Ah to 300 Ah.

For example, A 60-Ah battery, used in smaller automobiles, can supply a load current of 60/8 or 7.5 amperes for 8 hours without the cell voltages dropping below 1.7 volts. However this battery can supply less current for longer time (5 amps for 12 hours) or more current for a shorter time (60 amps for 1 hour).

Effect of temperature on Ah capacity of Lead-acid batteries

As in the case of primary cells, the capacity of a lead-acid cells also decreases significantly with temperature. These cells lose approximately 0.75% of its rated ampere-hour (Ah) capacity for every 1°F decrease in temperature. At 0°F (-18°C), its capacity is only 60% of the value at 60°F (15.6°C). In cold weather, therefore, it is very important to have an automobile battery always fully charged. In addition, at very cold temperature, the electrolyte freezes more easily as it is diluted by water in the discharged condition.

TIP: Keep the batteries always fully charged especially in cold weather conditions.

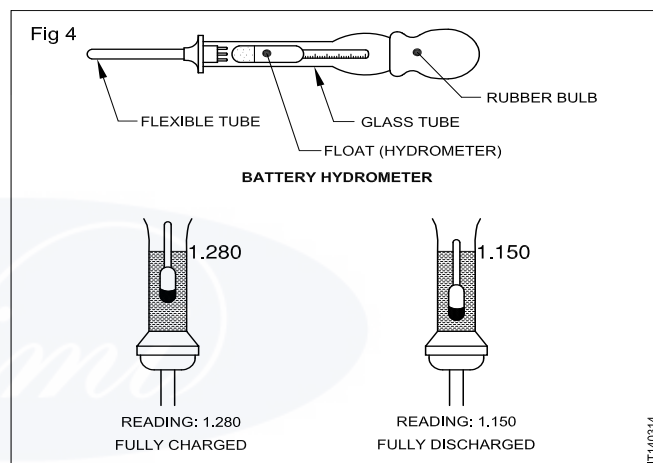
Specific gravity of electrolyte

Specific gravity is a ratio comparing the weight of a substance with the weight of water. The specific gravity of water is taken as 1 as a reference. For instance, specific gravity of concentrated sulphuric acid is 1.835.

This means, sulphuric acid is 1.835 times heavier than water for the same volume.

In a fully charged lead-acid cell, the specific gravity of the electrolyte, which is a mixture of sulphuric acid and water should be 1.28 at room temperature of 70 to 80°F. As the cell discharges, more and more water gets released into the electrolyte, lowering the specific gravity. When the specific gravity of the electrolyte falls down to about 1.150, the cell can be taken as fully discharged. Hence, the state of discharge of a lead-acid cell can be found out by measuring the specific gravity of its electrolyte.

The specific gravity of electrolyte is measured using a instrument known as *Battery hydrometer* as shown in Fig 4 below.



The importance of specific gravity can be seen from the fact that the open circuit voltage(V) of lead-acid battery is approximately given by,

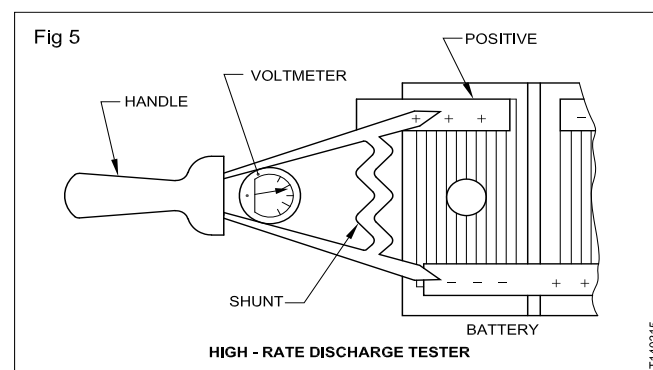
$$V = \text{Specific gravity} + 0.84.$$

For instance, if the specific gravity is 1.280 then,

$$V = 1.280 + 0.84 = 2.12V$$

Instrument for testing condition of cells - High rate discharge tester

The internal condition of a lead-acid battery cell is determined by this test. A low range (0-3V) voltmeter is shunted by a low resistance as shown in Fig 5.



The two terminal prods are pressed on to the terminals of a cell for testing. For fully charged cell the meter pointer points in the range of *full charge* on the meter scale. A sulphated old cell will show the discharge reading. The meter is having three colours red, yellow and green; red for fully discharged, yellow for half charged and green for fully charged condition of the cell respectively.

Topping up of lead-acid battery cells

In normal working condition of a lead-acid battery, the level of the electrolyte solution should be such that all the plates of the cells are fully immersed. If the level of the electrolyte is found to be less, then distilled water should be added to the indicated level of the cell through the vent plugs. This process of maintaining the level of electrolyte in lead-acid battery cell is called topping up.

Do not add tap water or well water for topping up. This will reduce cell life.

When a lead acid battery is being charged, the vent plugs are to be kept open for the gas produced to escape freely into air.

In case of lead-acid batteries used as back-up DC supply in un-interrupted power supplies (UPS), since charging and discharging of batteries is a continuous process, the vent plugs of the batteries will have several holes made on it for the gases produced during charging.

Un-interrupted power supplies are used in Hospitals, Computers etc., where the power failure may prove very costly. Details of UPS is discussed in detail in further units.

Care and routine maintenance of lead-acid batteries

- DO NOT use battery if it is discharged beyond the minimum value of 1.7V per cell.
- DO NOT leave a discharged battery in that condition for a long time. Even if not in use, keep the battery always fully charged.
- Always maintain the level of the electrolyte 10 to 15 mm above the top of the plates by adding suitable quantity of distilled water (NOT tap water).
- DO NOT add sulphuric acid to maintain specific gravity.
- Keep the vent openings in the filling plug always open to prevent build-up of high pressure due to the gases formed. At least the vent plug should have holes made in it.
- Wash off the acid and corrosion on the battery top using moist cloth, baking soda and water.
- Clean the battery terminals and metal supports up to the bare metal and apply vaseline or petroleum jelly over its surface.
- DO NOT test a discharged battery using a 'High rate discharge tester'.

Table in appendix D gives a list of the commonly used primary and secondary batteries and their details.

Recharging lead-acid batteries

Recall that lead-acid batteries are rechargeable. Once the cell voltages of a lead-acid battery falls below 1.8 V, the battery needs recharging. This discharged state of battery can be found by measuring the specific gravity of the electrolyte (1.150) or by measuring the voltage across the cells of the battery.

To charge a lead-acid battery, an equipment known as *Battery charger* is used. A battery charger is nothing but a dc voltage source which can supply the necessary voltage and charging current to the battery.

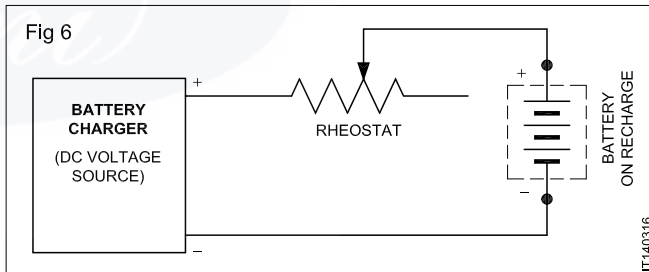
There are two main methods of charging batteries. They are;

- (1) Constant current battery charging
- (2) Constant voltage battery charging.

1 Constant current battery charging

In this method of charging batteries, the charging current supplied to the battery is kept at a prescribed (by the battery manufacturer) constant value. The amount of this constant current varies depending upon the Ah capacity of the battery. The value of constant charging current should not be excessive as this would cause excessive gassing. Excessive value current rises the cell temperature above the safe limit (generally 40°C) which will reduce the life of the battery.

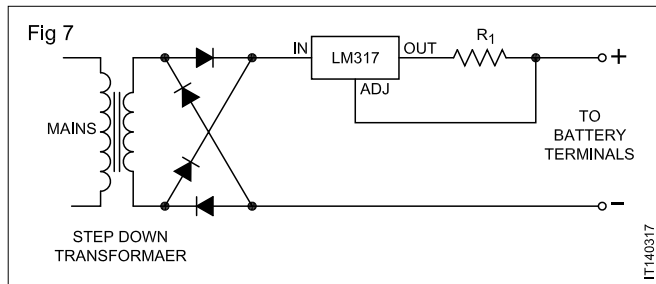
Fig 6 shows a very simple method of constant current charging system.



In constant current charging, the output dc voltage of the charger will be generally twice the nominal voltage of the battery to be charged. But, the charging current is controlled by varying the rheostat connected in series with the battery. For example, to charge a 12 V battery, the dc voltage source can be 24 V, but the charging current will be kept controlled say, 1 ampere with the help of the rheostat.

With the introduction of voltage regulator integrated circuits like LM317, it has become very simple and less expensive to make constant current battery chargers. Fig 2 shows a simple constant current battery charger using LM317. This charger can be used for any type of battery charging as long as the charging current is less than 1.5 Amperes.

In Fig 7, the resistor R_1 between the adjustment terminal and the output of the regulator sets a constant output current given by,



and Wattage of $R_1 \geq (I_{out})^2 \cdot R_1$

$$I_{out} = \frac{1.25}{R_1}$$

Current can be set at any value between 10 mA and 1.5 A in the circuit at Fig 7. To have higher currents, suitable external power transistors can be used. In Fig 7, the input voltage to the regulator IC (LM317) should be 1.5 times the battery voltage (to be charged) plus 3 V. LM317 used in Fig 7 is immune to output shorts or reverse battery connections. Hence, the charger will always be safe.

Refer application notes on LM317 regulators for improved circuits of constant current battery chargers.

The disadvantage of constant current battery charging is that it takes comparatively long time to fully charge the battery. But, the charge efficiency, which is defined as,

$$\text{Charge efficiency} = \frac{\text{Charge stored by the battery}}{\text{Charge supplied to the battery}}$$

is high compared to constant voltage battery charging.

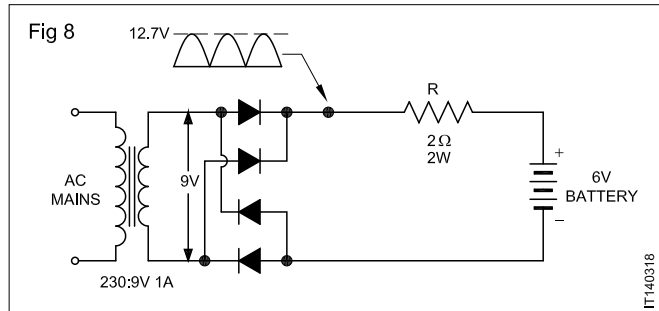
2 Constant voltage battery charging

In this method, the voltage applied across the battery terminals is kept constant, but no control is imposed on the charging current. Therefore, the battery draws large charging current in the beginning and as the cells get charged, the charging current decreases to a small value.

In this method, the time required for charging is reduced to half compared to the constant current charging. But, the charge efficiency gets reduced by approximately 10%.

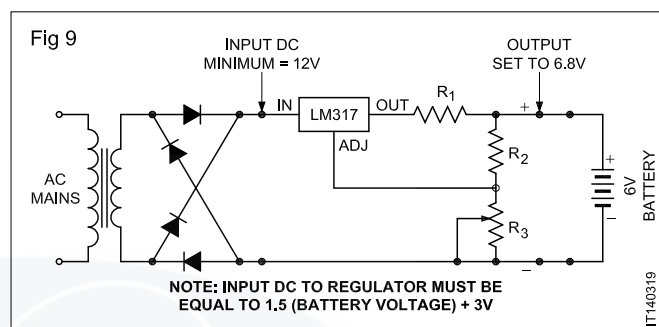
In constant voltage charging, the voltage applied to the cells for charging should be fixed at about 2.3 to 2.5 volt per cell and not more. For instance, for a 12 volts car battery, the dc voltage output of the charger should be between 14 V to 15 V.

A very simple constant voltage battery charger is shown in Fig 8. For reducing the cost of the circuit, filter capacitors are not used at the output of the bridge rectifier. Because of the absence of filter capacitors the battery gets charged only during the peaks of the rectified pulses. Therefore, the peak output voltage of pulsating dc should be at least two times the battery voltage in which fullwave/bridge rectifier is used or four times if half wave rectifier is used.



Resistor R is used to limit the initial charging surge current from becoming excessively high. This is because excessive current may damage the diode and transformer of the battery charger unit.

Innumerable number of battery charger circuits are available for constant voltage charging. The 3-pin adjustable voltage regulators such as LM317 can be used as a constant voltage battery charger as shown in Fig 9.



The advantages of using LM317 in the battery charger circuit are;

- protects the charger even if the battery is connected backwards.
- Protects the charger in case of output short.
- since LM317 is an adjustable voltage regulator (1.2V to 37V) a single circuit can be used to charge from even a single cell to a battery. The output voltage is easily adjustable with the help of resistors R_2 and R_3 (Fig 4) and is given by,

$$V_{out} = 1.25 + \left(1 + \frac{R_3}{R_2} \right)$$

In Fig 9, resistor R1 is used to limit the peak charging current.

NOTE: For further details refer LM317 application note.

TRICKLE Charging

Whenever a storage battery is used as an emergency reserve, as in the case of un-interrupted power supply (UPS), it is necessary to keep the batteries fully charged and ready for use at any time the mains supply fails.

A fully charged battery, which is not connected to any load is expected to maintain its terminal voltage. But, due to internal leakage in the battery and other open circuit losses, the battery voltage slowly falls even in idle or open circuit condition. Therefore, to keep it in fully charged condition, the battery should be supplied with a charging current which is small and just sufficient to compensate the idle condition or open circuit losses. This small current charging is known as *Trickle charging*. Trickle charging keeps the battery always fully charged and in ready to use condition, so that, the battery can be fully made use of in emergency conditions.

For example, if the normal charging current for a battery is 4 amps, then the trickle charge current would be around 20 to 40 mA (1:200 or less).

FLOAT Charging

In this method, the charger and the battery are always connected in parallel for supplying current to the load. The charger provides current for the load and the current necessary to keep the battery fully charged. The battery here is an auxiliary source of dc power.

Note that an automobile battery is in a floating-charge circuit. The battery charger is an ac generator or alternator with rectifier diodes driven by a belt from the engine. When the car is started, the battery supplies the cranking power. Once the engine is running, the alternator charges the battery.

Precautions for charging battery

- Batteries should never be charged or discharged at a higher rate than that specified by the manufacturer. This will weaken the plate structure and reduce battery life.
- Batteries should be recharged as early as possible after a discharge.

- The *high rate discharge tester* should be used only on fully charged batteries and that too for less than ten seconds.
- The specific gravity of the electrolyte should be checked before and after a battery is charged.
- The battery charging room should always be well ventilated for the gases to escape freely.
- If a battery is not being used for a long period, then the battery should be put on trickle charge.
- The vent plugs should be kept open while charging, for free liberation of gases.
- Avoid overcharging and discharging at a high rate. This causes the plates to bend from their position and buckle.
- Make sure that, while charging, the positive terminal of the charger is connected to the positive terminal of the battery, and the negative terminal of the charger to the negative terminal of the battery. Wrong connection causes very high current which can seriously damage both the battery and the charging unit.
- Make sure the cell temperature during charge does not exceed the limit specified (43°C) as per the manufacturer's instruction.
- The rate of charging at the end of the period called finish rate is most important. It must not exceed the value recommended by the manufacturer.
- During recharging, lead acid batteries produce flammable gases. An accidental spark can ignite these gases, causing an explosion inside the battery. Such an explosion can break the battery case and throw acid on the people and equipment in the area.
- Always wear safety glasses when working with lead acid cells and batteries. If acid does come in contact with clothing or with the skin, immediately flush with clean water. Wash your hands in soap and water after handling batteries.

Space for recording additional information



Characteristics & specification of Diodes

Objectives: At the end of this lesson you shall be able to

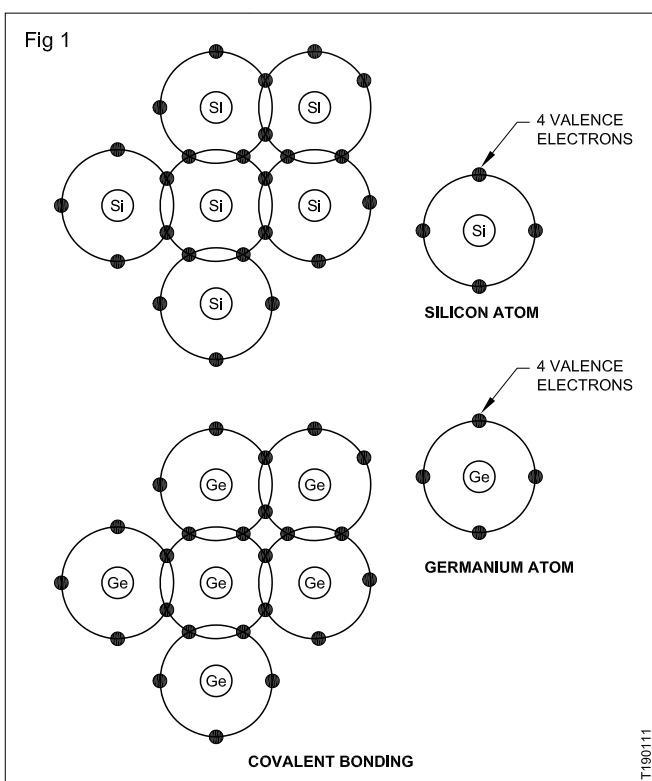
- state the meaning of intrinsic and extrinsic semiconductors
- define the P and N type semiconductor material
- state the meaning of donors and acceptors in doping
- state the unique property of a PN junction forward and reverse biasing
- list the different classifications of diodes
- state how to test a diode using an ohmmeter
- list the different types of diodes packaging
- state the different industry standards for diode type numbering scheme
- list the important specifications of a diode
- refer the diode data manual and find the specifications and equivalent of given diode
- explain the meaning of cut-in or knee or threshold voltage
- explain the meaning of forward and reverse biasing of diode
- explain the effect of temperature in diodes.

SEMICONDUCTORS

Semiconductors are materials whose electrical property lies between that of Conductors and Insulators. Because of this fact, these materials are termed as semiconductors. In conductors the valence electrons are always free. In an insulator the valence electrons are always bound. Whereas in a semiconductor the valence electrons are normally bound but can be set free by supplying a small amount of energy.

Semiconductor theory

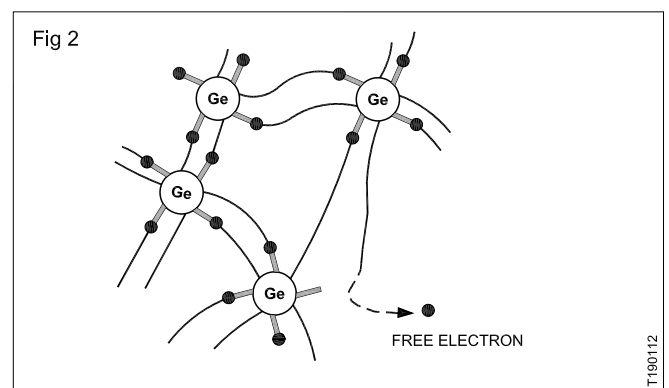
Basic semiconductor materials like other materials have crystal structure. The atoms of this structure, are bonded to each other as shown in Fig 1. This bonding is known as covalent bonding.



Intrinsic semiconductors

The most important of the several semiconductor materials are Silicon (Si) and Germanium (Ge). Both these semiconductor materials have four valence electrons per atom as shown in Fig 1. These valence electrons, unlike in conductors, are not normally free to move. Hence, semiconductors in their pure form, known as Intrinsic semiconductors, behave as insulators.

However, the valence electrons of a semiconductor can be set free by applying external energy. This energy will tear-off the bound electrons from their bond and make them available as free electrons as shown in Fig 2. The simplest method of turning bonded valence electrons into free electrons is by heating the semiconductor.



The higher the temperature to which the semiconductor is heated, more the bound electrons becoming free and will be able to conduct electric current. This type of conduction in an intrinsic semiconductor (pure semiconductor) as a result of heating is called intrinsic conduction.

From the above said phenomena, it is important to note that semiconductors are temperature-sensitive materials.

Extrinsic semiconductor

The number of free electrons set free by heating a pure semiconductor is comparatively small to be used for any useful purpose. It is found experimentally that, when a small quantity of some other materials such as Arsenic, Indium, Gallium etc. is added to pure conductor material, more number of electrons become free in the mixed material. This makes the semiconductor have higher conductivity.

These foreign materials added to the pure semiconductor are referred to as impurity materials.

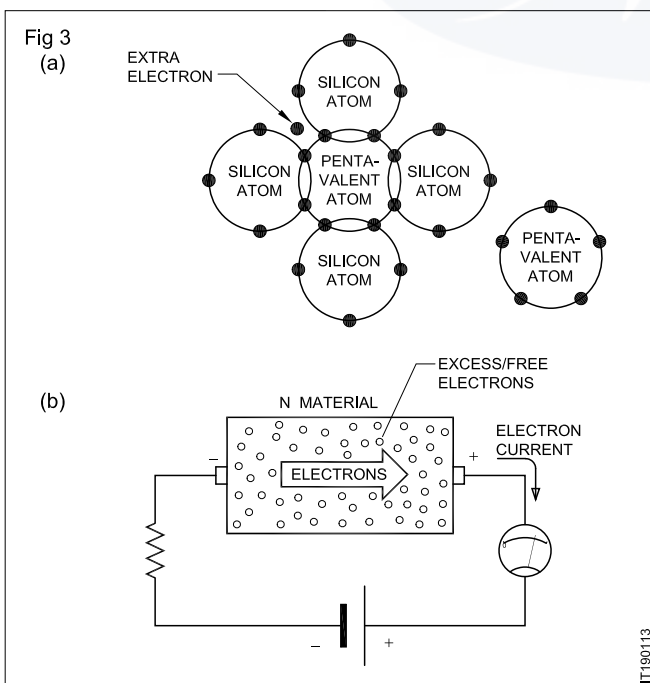
The process of adding impurity to an Intrinsic semiconductor material is known as Doping. Since the doped semiconductor materials are no longer pure, they are called impure or **extrinsic semiconductors**.

Depending upon the type of impurity used, extrinsic semiconductors can be classified into two types;

1 N-type semiconductors

When a pentavalent material like Arsenic (As) is added to a pure Germanium or pure Silicon crystal, one free electron results per bond as shown in Fig 3a. As every arsenic atom donates one free electron, arsenic is called the donor impurity. Since a free electron is available and since the electron is of a Negative charge, the material so formed by mixing is known as **N type material**.

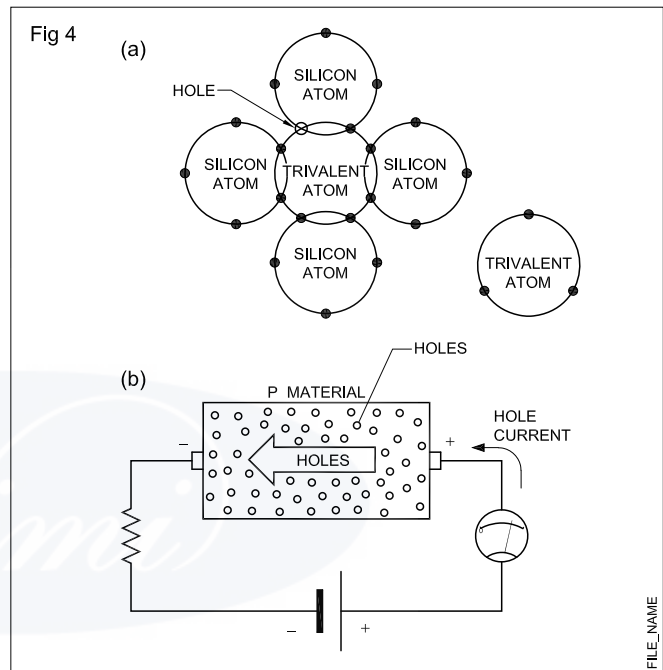
When a N-type material is connected across a battery, as shown in Fig 3b, current flows due to the availability of free electrons. As this current is due to the flow of free electrons, the current is called electron current.



2 P-type semiconductors

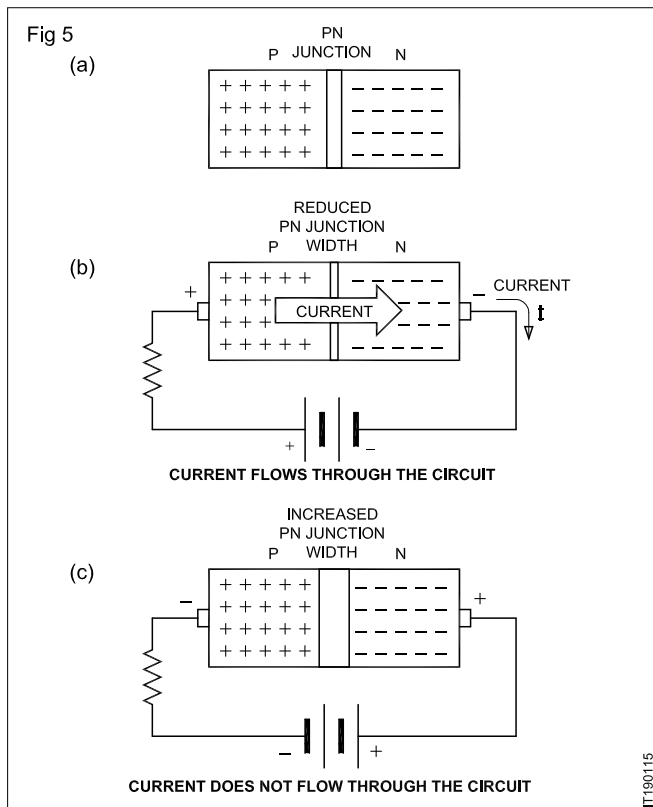
When a trivalent material like Gallium(Ga) is added to a pure Germanium or pure Silicon crystal, one vacancy or deficit of electron results per bond as shown in Fig 4a. As every gallium atom creates one deficit of electron or hole, the material is ready to accept electrons when supplied. Hence gallium is called acceptor impurity. Since vacancy for an electron is available, and as this vacancy is a hole which is of Positive charge, the material so formed is known as **P-type material**.

When a P-type material is connected across a battery as shown in Fig 4b, current flows due to the availability of free holes. As this current is due to flow of holes, the current is called hole current.

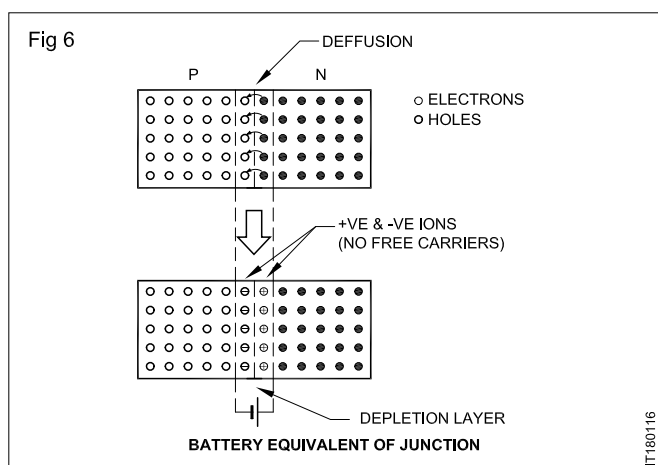


P-N junction

When a P-type and a N-type semiconductors are joined, a contact surface between the two materials called PN-junction is formed. This junction has a unique characteristic. This junction, has the ability to pass current in one direction and stop current flow in the other direction as shown in Fig 5. To make use of this unique property of the PN junction, two terminals one on the P side and the other on the N side are attached. Such a PN junction with terminals attached is called a **Diode**. The typical symbol of a PN-junction diode is shown in Fig 9a.

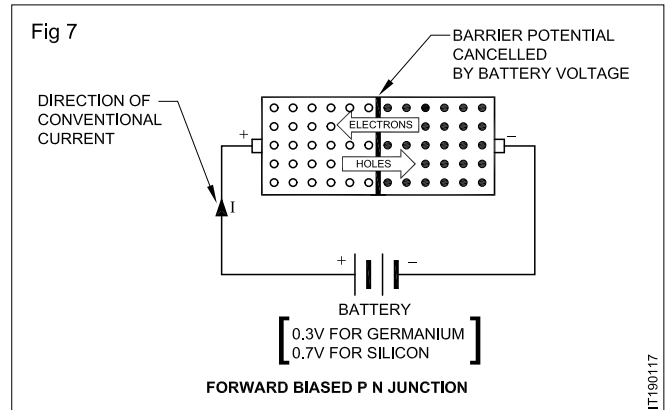


When a P and N material is put together, at the junction of P and N materials, as shown in Fig 6, some electrons from the N-material jump across the boundary and recombine with the hole near the boundary of the P-material. This process is called diffusion. This recombination makes atoms near the junction of the P-material gaining electrons and become negative ions, and the atoms near the junction of the N-material, after losing electrons, become positive ions. The layers of negative and positive ions so formed behave like a small battery. This layer is called the depletion layer because there are neither free electrons nor holes present (depleted of free carriers). This depletion region prevents further the movement of electrons from the N-material to the P material, and thus an equilibrium is reached.

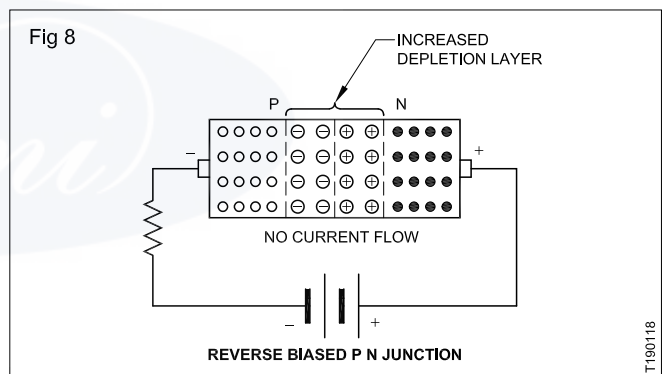


The internal voltage set up due to +ve and -ve ions at the junction is called barrier potential. If any more electrons have to go over from the N side to the P side, they have to overcome this barrier potential. This means, only when the electrons on the N side are supplied with energy to overcome the barrier potential, they can go over to the P side.

In terms of voltage applied across the terminals of the PN junction diode, a potential difference of 0.7V is required across the terminals in the case of silicon diode and 0.3V in the case of Germanium diode for the electrons, in order to cancel off the barrier potential and cross over the barrier as shown in Fig 7. Once the barrier potential gets canceled due to external voltage application, current flows through the junction freely. In this condition the diode is said to be forward biased.



When the applied external battery polarity is as shown in Fig 8, instead of canceling the barrier potential, the external battery voltage adds up to the barrier potential, and, hence, no current flows through the junction. In this condition the diode is said to be reverse biased.



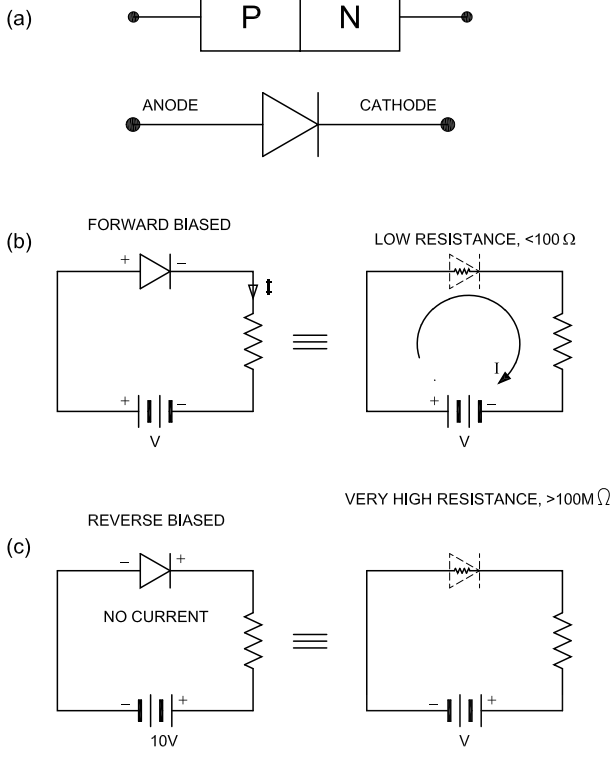
Since current flows through a PN junction diode when it is forward biased and does not when reverse biased, the diode can be thought of to be a unidirectional current switch.

Fig 9a shows the symbolic representation of a Diode. The two leads connected to the P and N terminals are known as Cathode and Anode.

To forward-bias a diode, the Anode should be connected to the +ve terminal of the battery and the Cathode to the -ve terminal of the battery, as shown in Fig 9b. When a diode is in the forward biased condition, the resistance between the terminals will be of the order of a few ohms to a few tens of ohms. Hence, current flows freely when a diode is forward biased.

On the other hand, when a diode is reverse biased as shown in Fig 9c, the resistance between the terminals will be very high, of the order of several tens of megohms. Hence, current does not flow when a diode is reverse biased. As a rule, the ratio of resistance in forward to reverse bias should be of at the minimum order of 1:1000.

Fig 9



Types of diodes

The PN junction diodes discussed so far are commonly referred to as rectifier diodes. This is because these diodes are used mostly in the application of rectifying AC to DC.

Classification of Diodes

1 Based on their current carrying capacity/power handling capacity, diodes can be classified as

- **low power diodes**
can handle power of the order of several milliwatts only
- **medium power diodes**
can handle power of the order of several watts only
- **high power diodes**
can handle power of the order of several 100's of watts.

2 Based on their principal application, diodes can be classified as,

- **Signal diodes**
low power diodes used in communication circuits such as radio receivers etc. for signal detection and mixing
- **Switching diodes**
low power diodes used in switching circuits such as digital electronics etc. for fast switching ON/OFF of circuits
- **Rectifier diodes**

medium to high power used in power supplies for electronic circuits for converting AC voltage to DC.

3 Based on the manufacturing techniques used, diodes can be classified as,

- **Point contact diodes**
a metal needle connected with pressure on to a small germanium(Ge) or silicon(Si) tip. See Chart 1 at the end of this lesson.
- **Junction diodes**
made by alloying or growing or diffusing P and N materials on a semiconductor substrate.

NOTE: Refer to Chart 1 at the end of this lesson for illustration of some of the above types of diodes.

Types of diode packaging

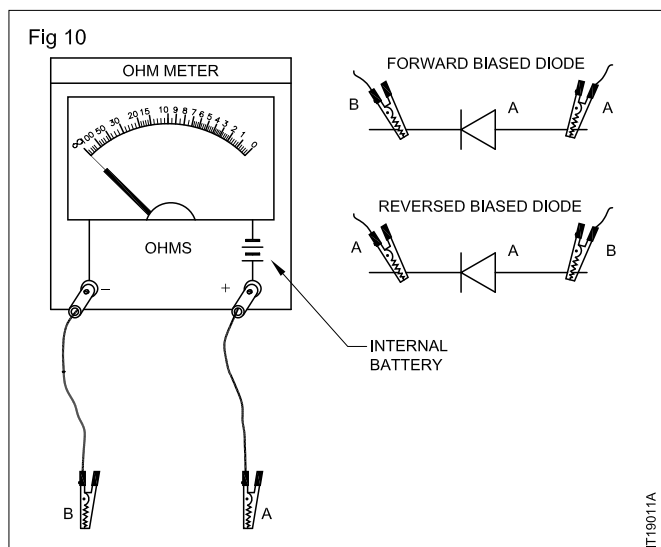
The type of packaging given to diodes is mainly based on the current carrying capacity of the diode. Low power diodes have either glass or plastic packaging. Medium power diodes have either plastic or metal can packaging. High power diodes will invariably have either metal can or ceramic packaging. High power diodes are generally of stud-mounting type. Refer illustrations given in Chart 1 at the end of this lesson, for illustrations of diode types.

Testing rectifier diodes using ohmmeter

A simple ohmmeter can be used to quickly test the condition of diodes. In this testing method, the resistance of the diode in forward and reverse bias conditions is checked to confirm its condition.

Recall that there will be a battery inside an ohmmeter or a multimeter in the resistance range. This battery voltage comes in series with the leads of the meter terminals as shown in Fig 10. In Fig 10, the lead A is positive, lead B negative.

NOTE: If the polarity of the meter leads are not known at first, the polarity of the meter leads can be determined using a voltmeter across the ohm meter terminals.



If the positive lead of the ohmmeter, lead A in the Fig 10, is connected to the anode of a diode, and the negative (lead B) to the cathode, the diode will be forward-biased. Current will flow, and the meter will indicate low resistance.

On the other hand, if the meter leads are reversed, the diode will be reverse-biased. Very little current will flow because a good diode will have very high resistance when reverse biased, and the meter will indicate a very high resistance.

While doing the above test, if a diode shows a very low resistance in both the forward and reverse biased conditions, then, the diode under test must have got damaged or more specifically shorted. On the other hand, a diode is said to be open if the meter shows very high resistance both in the forward and reverse biased conditions.

Polarity marking on the diodes

The cathode end of a diode is usually marked by a circular band or by a dot or by plus (+) sign as shown in Chart 1. In some diodes the symbol of the diode, which itself indicates the polarities, is printed on the body of the diode.

Type number or diode code number

Unlike resistors, capacitors or inductors, the diodes do not have any value that can be printed or coded on its body. The other reason for this is, there are almost innumerable types of diodes with varied current handling and other specifications. Hence, instead of printing its specifications on its body, all diodes will have a type number printed on their body. This type number carries a set of specifications which can be found out by referring to a diode data manual. Diode data manuals give data of several thousands of diodes from different manufacturers. Some of the popular type numbers of diodes are

OAx _x ,	xx - from 70 to 95.	examples: OA79, OA85 etc.,
BYxxx,	xxx- from 100 onwards,	examples: BY127, BY128 etc.
DRxxx,	xxx- from 25 onwards.	examples: DR25, DR150 etc.,
1Nxxxx	examples: 1N917	1N4001, 1N4007 etc.

SEMICONDUCTOR DIODES

Semiconductor diodes are used for various applications. Some of the major areas of application are listed below.

- Modulation and demodulation in communication receivers.
- Switching high speed and digital circuits.
- Low power and high power rectification.
- As surge protectors in EM relay and other circuits.
- For clipping, clamping wave-forms.

For different applications, diodes of different current carrying capacity, different PIV capacity and so on are required. Therefore, manufacturers of diodes make diodes to cater to

varied applications with different specifications. Before using a diode for a particular application, it is a must to find out whether the voltage, current, and temperature characteristics of the given diode match the requirement or not.

Important specifications of a diodes

- The material the diode is made-of: This could be Silicon or Germanium or Selenium or any other semiconductor materials. This is important because the cut-in voltage depends upon the material the diode is made-of. For example, in Ge diodes the cut-in voltage is around 0.3 V, whereas in Si diodes the cut-in voltage is around 0.7V.
- Maximum safe reverse voltage denoted as V_R or V_i that can be applied across the diode. This is also known as peak-inverse-voltage or PIV. If a higher reverse voltage than the rated PIV is applied across the diode, it will become defective permanently.
- Maximum average forward current, I_F or I_f that a diode can allow to flow through it without getting damaged.
- Forward voltage drop, V_F or V_f that appears across the diode when the maximum average current, I_F flows through it continuously.
- Maximum reverse current, I_{vr} that flows through the diode when the Maximum reverse voltage, PIV is applied.
- Maximum forward surge current, I_s that can flow through the diode for a defined short period of time.
- The maximum junction temperature in degree centigrade the diode junction can withstand without malfunctioning or getting damaged.
- Suggested application indicates application for which the diode is designed and produced.

The above listed specifications go with all rectifier diodes. As all these specifications cannot be printed on the physically small size diodes, the diodes are printed with a type number instead. When this type number is referred to in the manufacturer's manual, the detailed specifications for a particular type number of the diode can be obtained.

There are hundreds and thousands of diode manufacturers all over the world. To bring standardization for the diodes and other components manufactured by different manufacturers, the manufacturers and standards associations have set certain international standards for the benefit of users of the components. The principal industry standard numbering systems are dealt with here:

(a) The JEDEC type code

The EIA in USA maintains a register of **1N**, **2N** types familiarly known as Jedec types, which have world wide acceptance.

1N is used as a prefix for semiconductors with one junction. For example all 1N components refer to diodes because diodes have one junction. Prefix 2N is used with components having two junctions.

(b) The PRO-ELECTRON type code.

The Association International Pro-electron in Europe maintains a register of Pro-electron types which have wide acceptance in Europe.

Components in the Pro-electron system have,

- i) two letter and numeral code for consumer devices (Example, BY127 and so on).
- ii) three letter and numeral code for industrial devices. (Example, ACY17 and so on).

The first letter in the pro-electron type code indicates the type of semiconductor material used in making the device. Example, device numbers starting with A are made of germanium. Refer to Appendix D for further details and examples.

The second and third letter indicate the applications of the component. Example, in the type code BY127, the second letter Y indicates that it is a rectifier diode. Refer to Appendix D for further details.

The numeral after the second or third letter is the code number of its detailed voltage, current and temperature specification.

(c) The JIS type code

In Japan, the JIS, (Japanese Industrial Standards) code is used. This system of component numbering is almost universal. In this system, all component numbers start with 2S, followed by a letter and several numbers. Example. 2SB364. The letters after the S has the following significance:

A = pnp hf

B = pnp lf

C = npn hf

D = npn lf

Some components will have type numbers which does not match with any of the above said international standards. Then, these type numbers are particular to the individual manufacturers. These codes are generally referred to as manufacturer's house code. However, these type numbers may conform to one or more of the international standards. Almost all standard diode data books lists popular manufacturers house codes.

[H.I: List a few manufacturers house codes relevant to your region by referring to any popular diode data manual. Give one data manual each to a batch of 4 trainees. Write a few diode type codes on the black/white board and ask the batches to find the specifications.]

Diode equivalents

There are several occasions, especially while servicing electronic circuits, it may not be possible to get a replacement for a diode of a particular type number. In such cases one can obtain a diode having specification closest to the one to be replaced. Such diodes are referred to as equivalents.

Example: In a circuit, diode 1N 4007 is found to be defective. If 1N4007 is not available in stock, then, instead

of 1N4007, BY127 can be used because BY127 is the equivalent for 1N4007.

Some data books give the lists of equivalents. Appendix D gives the equivalent of some of the popular diodes used in general electronic circuits.

[H.I: Write a few popular diode type numbers generally used in your institute and instruct trainees to find its equivalents.]

Behaviour of diode when FORWARD BIASED

Fig 11a shows a forward biased silicon PN junction diode using a variable DC supply. When the applied voltage is slowly increased starting from 0 volts as long as the voltage across the diode V_F is less than that of the depletion barrier potential (0.7 volts for Si diodes), no current or a negligible current flows through the diode, and, hence, through the circuit. This is shown in the graph at Fig 11b.

But once the voltage V_F across the diode becomes equal to or greater than the barrier potential 0.6 to 0.7V, there will be a canceling effect of the barrier potential. Hence, the free electrons from the N region get pushed away by the -ve battery terminal (remember like charges repel) and cross over the junction, pass through the P region and get attracted by the + terminal of the battery. This results in the electron current passing through the diode, and, hence, through the Load.

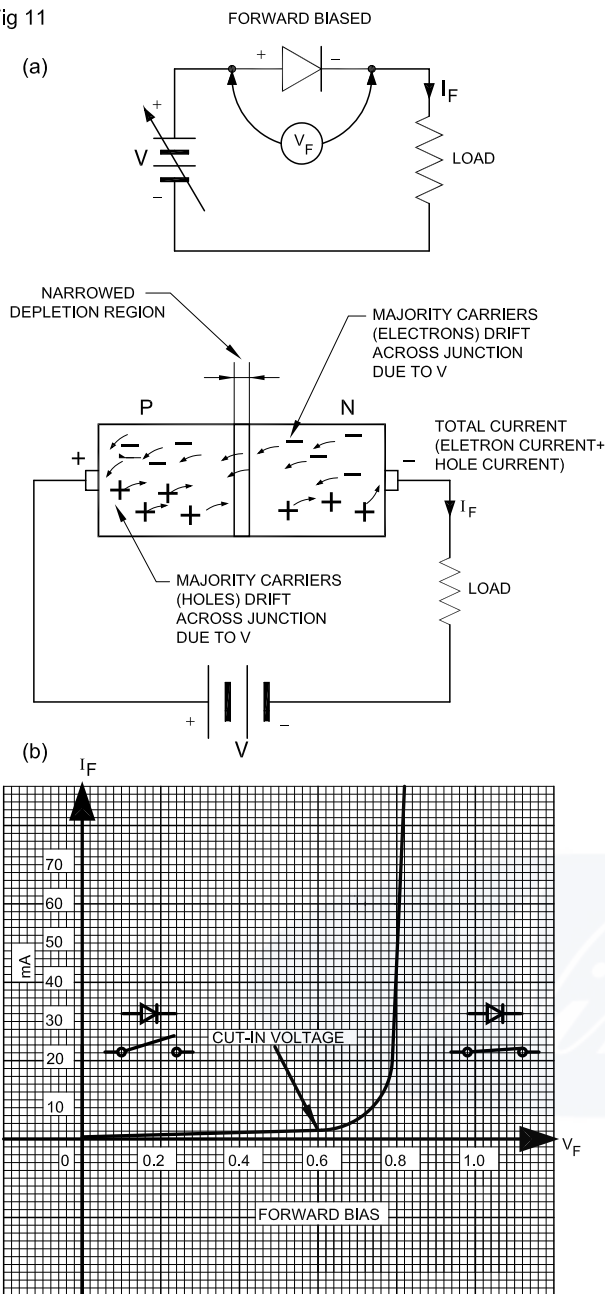
In a similar way, the holes in the P region are pushed away by the +ve battery terminal, resulting in hole current.

Thus current flows through the diode when the forward bias potential is higher than the barrier potential. This current flow through the diode is because of both electrons and holes. The total current in the circuit is the sum of the hole current and the electron current. Hence, diodes are called bipolar devices in which both hole current and electron current flows.

From the graph at Fig 11b, it can be seen that, once the forward voltage goes above 0.6V the diode starts conducting, resulting in considerable current through the circuit. This voltage level across the diode is referred to as cut-in or knee or threshold voltage.

If the applied forward voltage is further increased beyond the cut-in voltage, the depletion layer further narrows down allowing more and more current to flow through the diode. It can be seen from the graph at Fig 1b, that beyond the cut-in voltage, the current increases sharply for very small voltage increase across the diode. In this region, above the cut-in voltage, the forward biased diode behaves almost like a closed switch. The only limiting factor for the current at this stage is the maximum current the diode can handle without getting burnt or the junction getting punctured permanently. This current limit is given in diode data books as maximum forward current, I_{Fmax} .

Fig 11



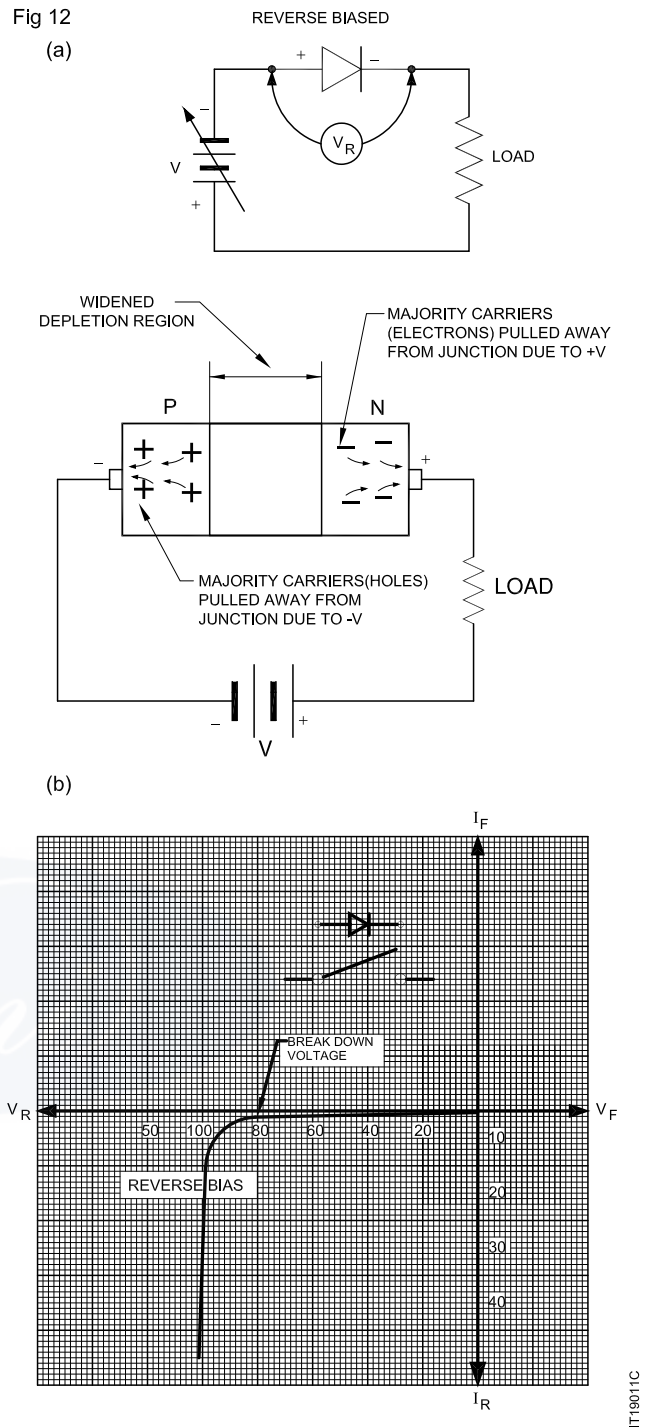
Behaviour of diode when REVERSE BIASED

When an external DC voltage is connected across the diode with the polarity as shown in Fig 12, the diode is said to be reverse biased.

In this condition, when the battery voltage is increased from 0 to several tens of volts, the polarity of the applied voltage instead of canceling the barrier potential, aids the barrier potential. This, instead of narrowing the depletion layer, widens the depletion layer. The widening of the depletion layer results in, not allowing the current to flow through the junction, and, hence, the load.

Referring to the graph shown in Fig 12b, it can be seen that there is no current even when the voltage V_R across the diode is several tens of volts.

Fig 12

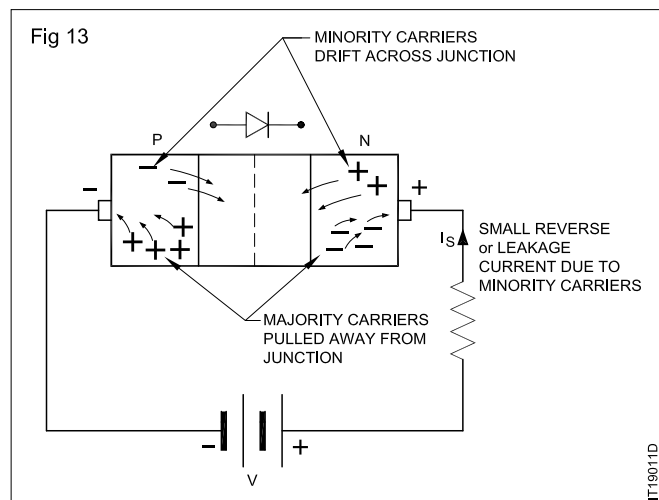


If the applied reverse voltage is kept on increased, say to hundred volts (this depends from diode to diode), at one stage the applied voltage V_R across the junction is so large that it punctures the junction damaging the diode. This results in shorting of the diode. This short results in uncontrolled heavy current flow through the diode as shown in graph at Fig 12b. This voltage at which the diode breaks down is referred to as reverse break-down or avalanche breakdown.

The maximum reverse voltage that a diode can withstand varies from diode to diode. This reverse voltage withstanding capability of a diode is referred to as the peak-inverse-voltage or PIV of the diode. This value for diodes is given in the diode data manual. The PIV of diodes varies from a minimum of 50 volts in small signal diodes to several thousands of volts in high power diodes.

Minority current in DIODES

When a PN junction is reverse biased, due to the increased width of the depletion layer, there can be no current through the diode. But, in practice there will be a small current of the order of a few nano-amperes or a few micro-amperes through the diode as shown in Fig 13.



The reason for this small current is due to the creation of a very limited number of free electrons and holes on both sides of the junction due to thermal energy.

Semiconductors are highly sensitive to temperature. Even a temperature of 25°C is sufficient to create a small number of electrons and holes resulting in a current of the order of a few nano-amperes. These current carriers created are referred to as minority current carriers. This current, due to minority current carriers, which flows through the diode when reverse biased, is known as reverse current or leakage current or saturation current, i_s . Based on experiments, for all silicon diodes, this reverse current doubles for each 10°C rise in temperature. For example, if i_s is 5nA at 25°C , it will be approximately 10nA at 35°C and so on.

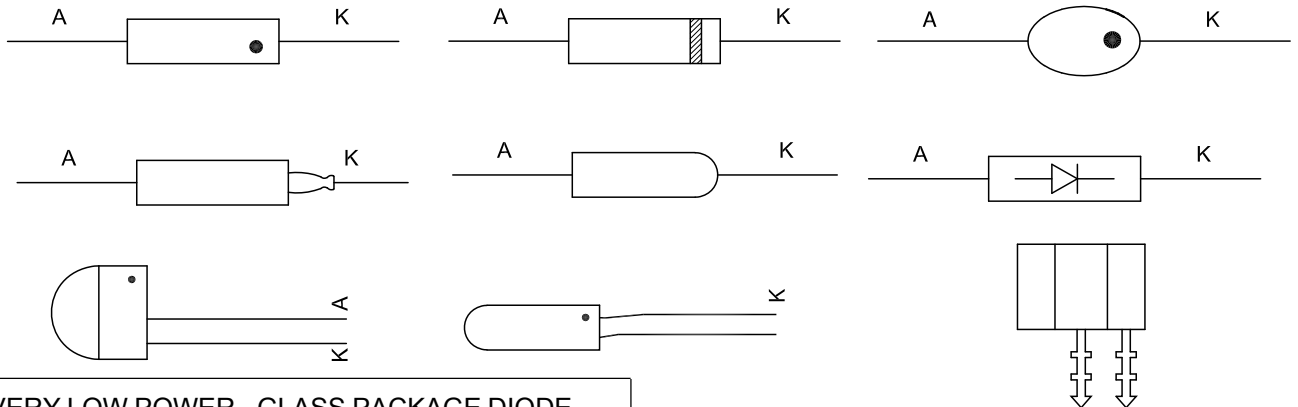
Effect of temperature on barrier voltage

It is known that semiconductors are highly sensitive to temperature. Since the functioning of a diode is basically due to the unique property of its junction and its barrier voltage, the barrier voltage also depends on the junction temperature. If the temperature of the junction is increased beyond a limit (25°C), electrons are produced due to thermal agitation in the semiconductor crystal structure. These electrons, having sufficient energy, drift across the junction. This decreases the barrier voltage. It is experimentally found that the barrier voltage decreases by $2\text{mV}/^{\circ}\text{C}$ increase in temperature. This reduced barrier voltage allows more current through the junction. More current heats up the junction further, reducing the barrier voltage further. If this cumulative effect continues, the junction will get damaged making the junction no more useful. Therefore, diodes should not be allowed to go above a specified temperature. This maximum limit a diode junction can withstand safely is given in the diode manual as junction temperature, $T_{j\text{max}}$.

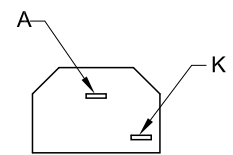
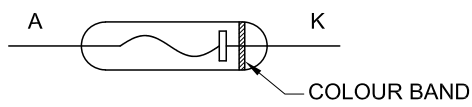
CHART - 1

TYPES OF DIODES AND PACKAGING

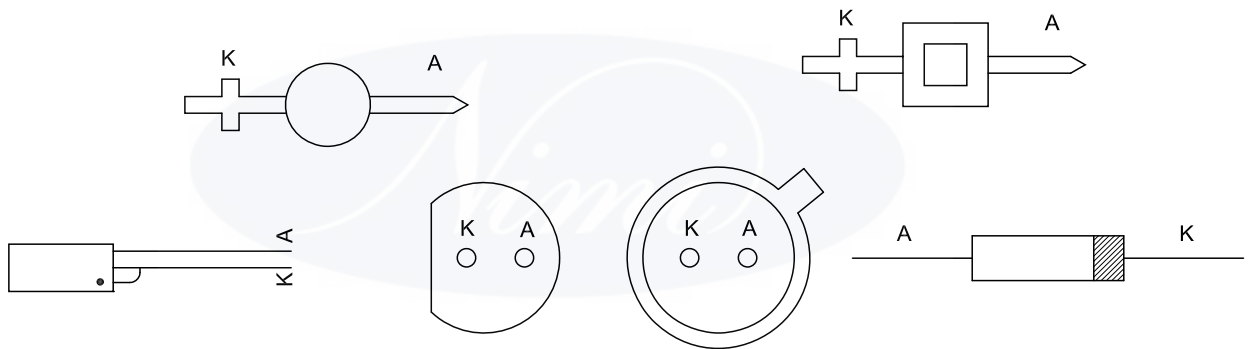
LOW POWER - PLASTIC PACKAGE DIODES



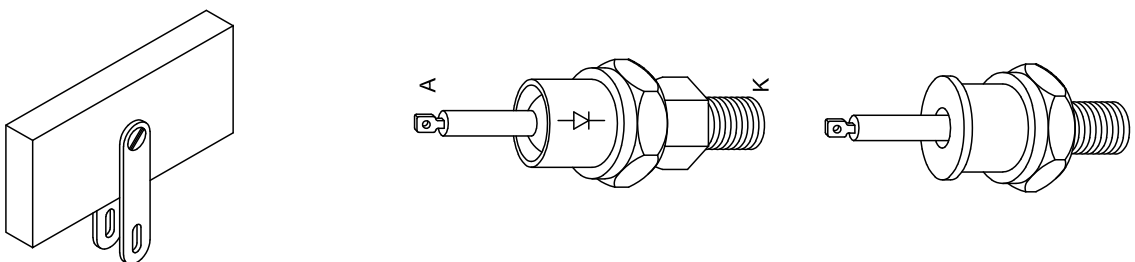
VERY LOW POWER - GLASS PACKAGE DIODE



LOW POWER - METAL PACKAGE DIODE



HIGH POWER - METAL PACKAGE DIODE



VERY HIGH POWER - METAL PACKAGE DIODES



Half wave & Full wave Rectifiers

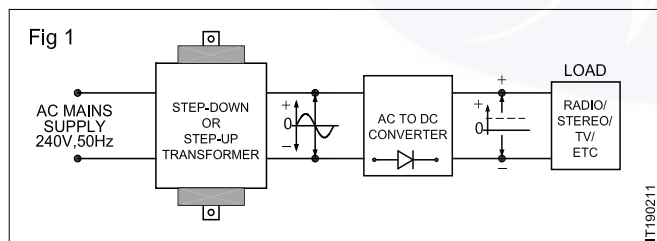
Objectives: At the end of this lesson you shall be able to

- state the need for converting AC to DC
- state the meaning of the term rectification
- describe the working of a half-wave and two diode full-wave rectifier
- state the output of a half-wave rectifier in terms of rms value of input AC, peak value of input AC
- state the output DC level and the current rating of diode in a two-diode fullwave rectifier
- state the ripple frequency in a half-wave rectifier and a full-wave rectifiers
- state the value of PIV in half-wave and full-wave rectifiers
- list the limitations of half-wave and a two diode full-wave rectifiers.

RECTIFICATION

Almost all electronic circuits need DC voltage for their working. This DC voltage can be obtained by dry cells and batteries. Use of a dry cell is practicable only in portable electronic circuits such as transistor radio, tape recorders etc. But in circuits requiring large voltages and currents, like high power audio amplifiers, television sets etc. batteries will not only be very expensive but also be voluminous.

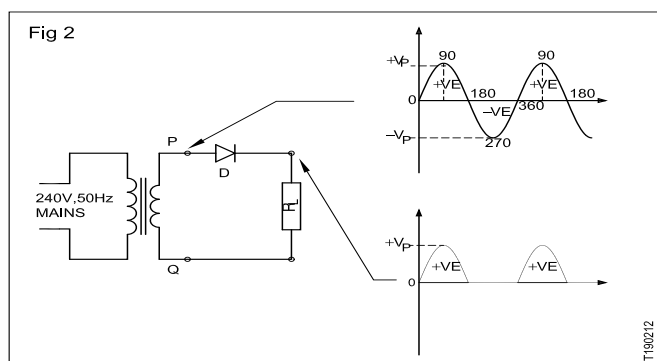
An alternative method of obtaining DC voltage is by converting the AC mains supply of 240V, 50Hz into DC voltage. This technique is not only convenient but also takes very small space compared to battery packs. This process of converting AC to DC is known as rectification. Fig 1 shows the principle of converting AC to DC of required voltage level.



The transformer will step-down or step-up the mains AC to the required level. The stepped-up or stepped-down AC from the output of the transformer is then converted to DC using diodes making use of their unique unidirectional property.

HALFWAVE RECTIFIER

The simplest form of AC to DC converter is by using one diode. such an AC to DC converter is known as half-wave rectifier as shown in Fig 2.



At the secondary of the transformer, across terminals P & Q, when seen on a CRO, the electric signal is a sinusoidal wave with its peak value of V_p and a frequency determined by the rate at which the alternations (+ve to -ve) are taking place. In Fig 2, the frequency is 50Hz as this voltage is taken from 50Hz AC mains supply.

If the voltage across P and Q is measured using an AC voltmeter, the voltmeter shows the rms (root means square) value, V_{rms} of the sinusoidal wave which will be less than the peak value. The relationship between V_{Peak} and V_{rms} is given by,

$$V_{rms} = 0.707 V_{peak} \quad \dots\dots[1]$$

conversely,

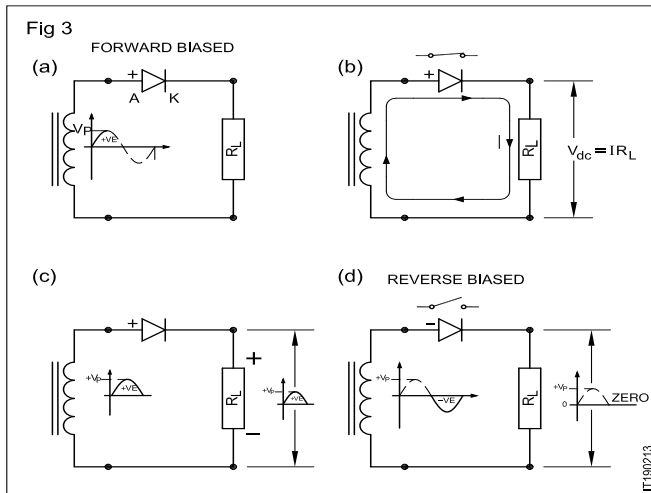
$$V_{peak} = \frac{V_{rms}}{0.707} = \sqrt{2} V_{rms}$$

When this sinusoidal signal is applied across the diode D as shown in Fig 2, the diode conducts (behaves as a closed switch) only during the +ve half cycle of the input sinusoidal voltage and does not conduct (behaves as an open switch) during the -ve half of the input sinusoidal voltage. This process repeats again and again thus producing a pulsating +ve wave-form at the output across the load, R_L as shown in Fig 2.

The operation of a half-wave rectifier circuit can be summarised with the help of Fig 3 as follows:

- 1) During the positive half cycle of AC input, the diode is forward biased as the anode of diode is positive as shown in Fig 3a.
- 2) Hence current flows from anode to cathode, through load R_L to secondary of transformer as shown in Fig 3b. The IR_L drop across load resistor R_L is the DC voltage V_{dc} with the polarity as shown in Fig 3b.
- 3) When the +ve half cycle of the input sinusoidal is completed, the voltage across the R_L will be a positive half sinewave as shown in Fig 3c. The peak of rectified voltage is also equal to the peak of the input AC voltage.
- 4) During the negative half cycle of the input AC, the diode is reverse biased as the anode of diode is negative as shown in Fig 3d.

- 5) Hence, the diode behaves as an open switch and no current flows through the load and hence there is no voltage output across load R_L as shown in Fig 3d.
- 6) After completing the $-ve$ half cycle, when the input signal goes positive again, the whole operation repeats starting from step 1.



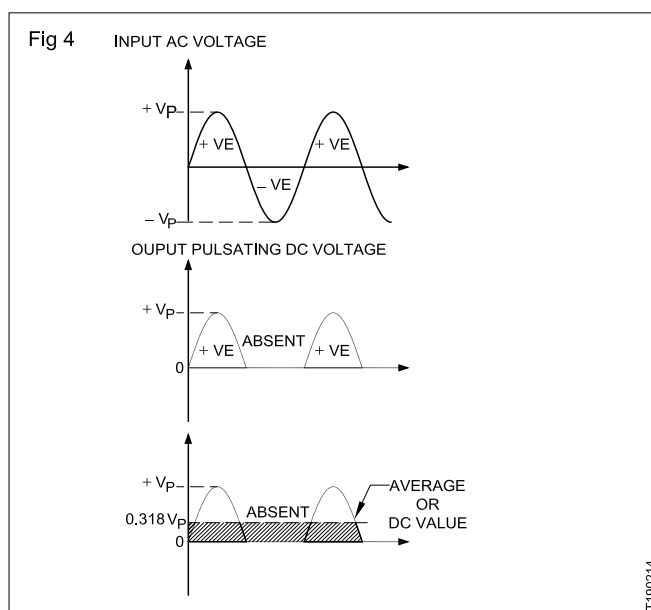
As can be seen from Fig 2, the output of the half-wave rectifier is always a $+ve$ voltage (DC) although it is pulsating. In other words, the output is either positive (during $+ve$ half cycle of the AC input) or zero (during $-ve$ half cycle of AC input) but never negative. Hence, the output of a rectifier is a pulsating $+ve$ DC voltage.

The circuit at Fig 2 is known as a half-wave rectifier as the rectification is done by the circuit only during one half cycle of the input AC signal.

Calculating output DC level in half-wave rectifiers

Two important points to note for calculating the output DC level of a halfwave rectifier are;

- the output of a halfwave (HW) rectifier across the load resistor is a pulsating DC whose peak voltage is equal to the peak value of the $+ve$ half cycle of the AC input as shown in Fig 4. This can be checked using an oscilloscope.



The small forward voltage drop of the diode (0.7 for Si) when forward biased is ignored for simplicity in understanding.

- the pulsating signal level is zero when the input AC is in the $-ve$ half cycle as shown in Fig 4.

Hence, When a DC voltmeter is connected across the load resistor R_L , the meter reads the average DC value of the pulsating signal. Ignoring the diode drop, the average DC value of the pulsating output in a half wave rectifier is given by,

$$V_{\text{average}} \text{ or } V_{\text{dc}} = 0.318 V_p \quad \dots\dots[2]$$

Example: If the total secondary voltage of the transformer (V_s) in Fig 2 is $24 V_{\text{rms}}$ (measured by AC meter), the output V_{dc} will be,

$$\text{From ...1, } V_{\text{peak}} = \frac{V_{\text{rms}}}{0.707}$$

$$\text{From ...2, } V_{\text{dc}} = (0.318) V_{\text{peak}} = \frac{V_{\text{rms}}}{0.707} = 0.45 V_{\text{S(rms)}}$$

Therefore, for a half-wave rectifier the level of output DC is given by,

$$V_{\text{dc}} = 0.45 V_{\text{S(rms)}} \quad \dots\dots[3]$$

Where $V_{\text{S(rms)}}$ is the input rms AC voltage.

In the example considered above, the output DC voltage at Fig 2 will be,

$$V_{\text{dc}} = 0.45 \times V_{\text{S(rms)}} = 0.45 \times 24 = 10.8 \text{ volts.}$$

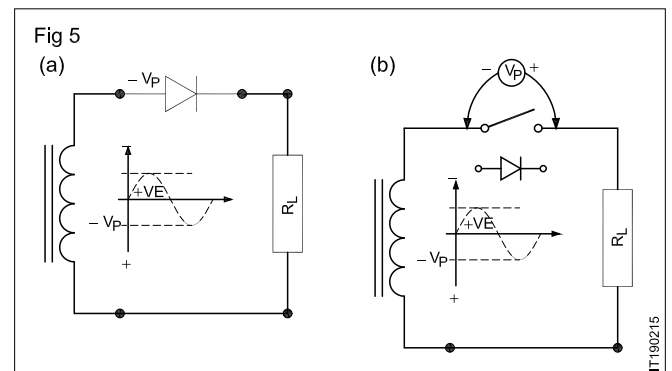
Ripple frequency

From Fig 4 it is evident that the frequency of the rectified pulsating DC is same as the frequency of the input AC signal. This is true for all half-wave rectifiers.

Peak inverse voltage

Fig 5a shows the half-wave rectifier at the instant the secondary voltage is at its maximum negative peak.

In this condition, since the diode is reverse biased, it behaves as an open switch as shown in Fig 5b. Since the diode is reverse biased, there is no voltage across the load R_L . Therefore, From Kirchhoff's voltage law, all the secondary voltage appears across the diode as shown in Fig 5b. This is the maximum reverse voltage that appears across the diode in the reverse biased condition.



This voltage is called the peak reverse voltage or more commonly as the **peak inverse voltage (PIV)**. Therefore, in a half-wave rectifier the peak inverse voltage across the diode is equal to the -ve peak value of the secondary voltage $V_{S(\text{peak})}$. Since, the -ve peak voltage and +ve peak voltage in a sinusoidal wave is same in magnitude, the peak inverse voltage (PIV) across the diode in a halfwave rectifier can be taken as a $V_{S(\text{peak})}$.

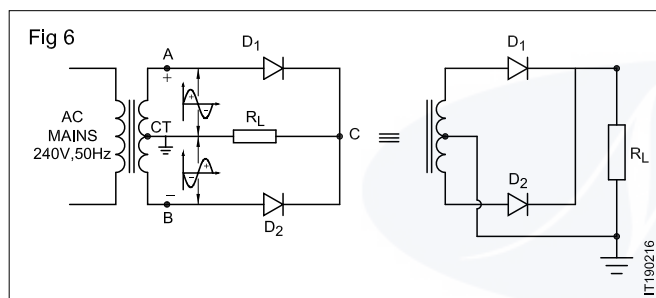
In the example considered earlier, the PIV across the diode will be,

$$V_{S(\text{peak})} = \frac{V_{S(\text{rms})}}{0.707} = \frac{24}{0.707} = 33.9 \approx 34 \text{ volts.}$$

To avoid break down of the diode used, the PIV appearing across the diode of the designed HW rectifier must be less than the PIV rating of the diode. For instance, in the above example to avoid break down of the diode, the PIV rating of the diode should be greater than 34 volts.

TWO-DIODE FULLWAVE RECTIFIER

In a half-wave rectifier there is no rectification action during the -ve half cycle of the input AC voltage. Because of this the output DC level is low ($0.318 V_{S(\text{peak})}$). This limitation of a half-wave rectifier can be overcome by using two diodes and a centre-tap-transformer as shown in Fig 6.



In Fig 6, each diode and the common load resistor R_L form two independent half-wave rectifiers. Because of the centre-tapped secondary winding, each diode receives only half of the total secondary voltage.

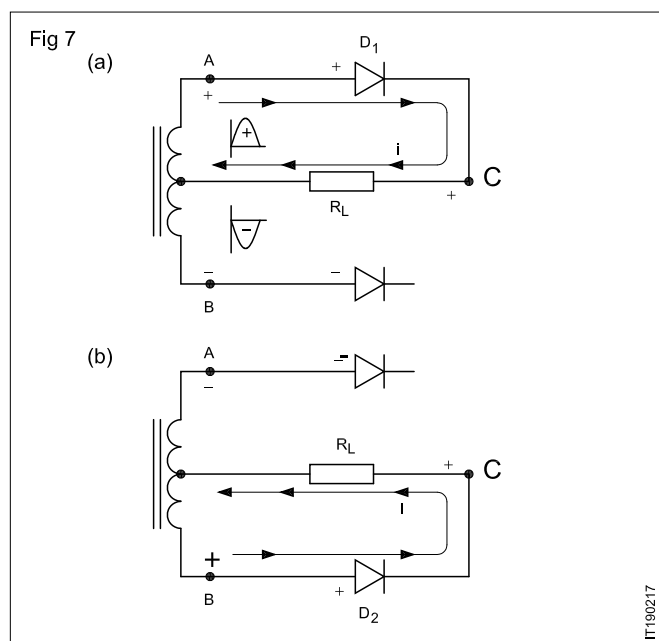
The opposite ends of the transformer secondary windings always have opposite polarities with respect to the centre tap. As shown in Fig 6, when end A of the secondary winding is positive, the bottom end B will be negative.

For the polarity shown in Fig 7a, the anode of D_1 is positive and, hence, forward biased. Whereas, the anode of D_2 is negative, and, hence, does not conduct. Current flows from the transformer(end A) $\rightarrow D_1 \rightarrow R_L \rightarrow$ back to the transformer centre-tap. This direction makes point C across the load R_L as the +ve terminal of the output DC voltage.

During the next half cycle, end B of transformer is +ve and A is -ve as shown in Fig 7b. Hence the anode of D_2 is +ve and this diode conducts whereas D_1 does not. Current flows from the transformer(end B) $\rightarrow D_2 \rightarrow R_L \rightarrow$ back to the transformer centre-tap. This direction of current again makes point C across the load R_L as the +ve terminal of the output DC voltage.

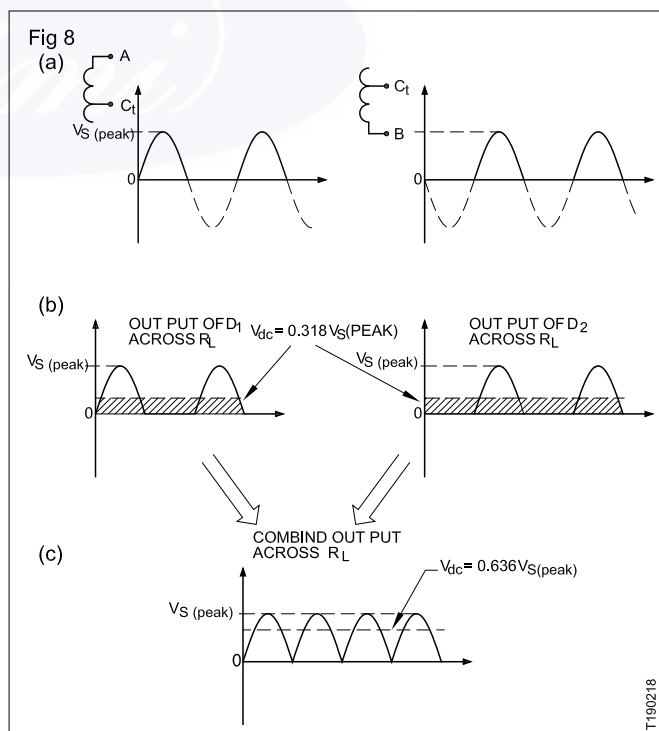
It is important to note the following two points;

- At any instant of time either D_1 or D_2 conducts but never both.



- While any of the two diodes is conducting, the rectified current i , flows through R_L in the same direction as shown in Fig 7a and Fig 7b. So the DC output voltage is positive at the common cathodes of the diodes D_1, D_2 .

Fig 8 shows the input and output wave-forms of a fullwave rectifier. As can be seen from Fig 8, the rectifier works both in the positive and negative half cycles of the AC input to produce a DC output across the load resistor R_L . Although only one diode conducts at a time, the outputs are



combined in R_L . Hence, fullwave rectifier provides double the DC current to the load compared to that of a half-wave rectifier.

Since both the half cycles of the input AC signal are rectified by the circuit at Fig 6, this circuit is known as a fullwave rectifier. Since this fullwave rectifier uses two diodes this circuit is also known as two diode fullwave rectifier.

Output DC level in a two-diode full wave rectifier

Since a full wave rectifier is nothing but a combination of two half-wave rectifiers, the average or DC value of a full wave rectifier is naturally twice the output of a half-wave rectifier driven by the same secondary voltage.

From Fig 8 it is evident that the average or DC value of a full wave rectified output is

$$V_{dc} = 0.318 V_{S(peak)} + 0.318 V_{S(peak)}$$

$$V_{dc} = 0.636 V_{S(peak)}$$

where, $V_{S(peak)}$ is the equal peak voltage between the centre-tap and any one end A or B of the transformer secondary.

In terms of $V_{S(rms)}$, V_{dc} of a full wave rectifier is given by,

$$V_{S(rms)} = 0.707 V_{S(peak)}$$

Therefore,
$$V_{dc} = 0.636 \cdot \frac{V_{S(rms)}}{0.707}$$

$$= 0.9 V_{S(rms)}$$

Example: Suppose the secondary voltage of the transformer is 24-0-24 $V_{(rms)}$, the DC output voltage of a full wave rectifier using this transformer will be,

For a two diode full wave rectifier

$$V_{dc} = 0.9 V_{S(rms)}$$

Therefore, in the given example,

$$V_{dc} = 0.9 \times V_{S(rms)} = 0.9 \times 24 = 21.6 \text{ volts.}$$

Ripple frequency in a full wave rectifier

From Fig 8c it can be seen that two cycles of output occur for each input cycle of AC voltage. This is because, the full wave rectifier has inverted the negative half-cycle of the input voltage. As a result, the output of a full wave rectifier has a frequency double the input AC frequency. If mains AC is used as input to a full wave rectifier, since mains frequency is 50 Hz, the output frequency of the pulsating DC will be 100 Hz.

This increased ripple frequency has certain advantages when the pulsating DC is smoothed. This will be dealt with in further lessons.

Peak inverse voltage

Fig 9 shows the full wave rectifier at the instant the secondary voltage reaches its maximum positive value.

Applying Kirchhoff's law around the outside loop, we get,

$$2V_{S(peak)} - \text{Reverse voltage (PIV) across } D_2$$

$$+ \text{Forward voltage across } D_1 = 0$$

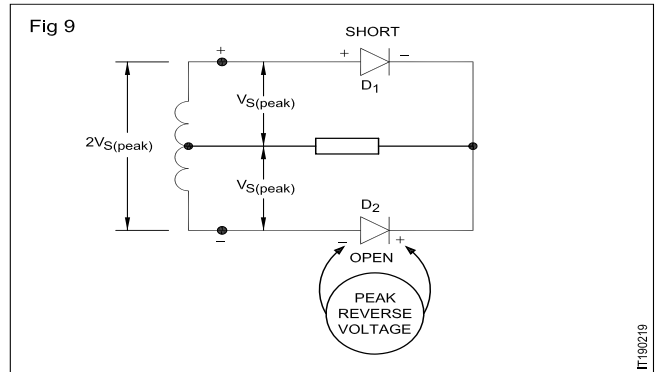
Neglecting the small forward voltage across D_1 we have,

$$2V_{S(peak)} - \text{PIV across } D_2 + 0 = 0$$

$$\text{or PIV across } D_2 = 2V_{S(peak)}$$

From the above it can be seen that each diode in a full wave rectifier must have a PIV rating greater than the peak value of the full secondary voltage. ($2V_{S(peak)}$)

In the example considered earlier, the PIV of diodes should be greater than 34V.



Current rating of diodes in a full wave rectifier

If the load, R_L connected in the full wave rectifier is, say 10W, the DC current through it will be,

$$I_{dc} = \frac{V_{dc}}{10\Omega}$$

In the example considered above, $V_{dc} = 21.6$ volts.

$$\text{Therefore, } I_{dc} = \frac{21.6}{10} = 2.16 \text{ amps.}$$

It is interesting to note this current I_{dc} is shared by the two diodes D_1 and D_2 . This is because each diode conducts only for one half cycle. Therefore, the DC current through each diode is half the total DC load current I_{dc} . Hence, the maximum current through each diode with 10W load will be $2.16/2 = 1.08$ amps. From this it follows that the current rating ($I_{f(max)}$) of each diode need only be half the maximum/rated load current.

NOTE: In a half wave rectifier since there is only one diode, the current rating of the diode used should be the maximum current through the load unlike in the case of a full wave rectifier in which the current rating of the diodes used is only half the maximum current through the load.

Example: In a two diode full wave rectifier, with a load current requirement of 1.8 amps, what should be the current ratings of the diodes used?

Since it is a two diode full wave rectifier, the current rating of each diode should be $\frac{1}{2}$ the total load current.

Therefore, $I_{f(max)}$ of diodes should be $\frac{1.8 \text{ amps}}{2} = 0.9$ amps.

It is fine if a diode of 1 amp current rating is used for this rectifier circuit.

Disadvantages of TWO DIODE full wave rectifier

The full wave rectifier using two diodes and centre tap transformer has the following disadvantages;

- A centre-tapped transformer that produces equal voltages on each half of the secondary winding is difficult to manufacture and, hence, expensive.
- Centre-tapped transformers are generally bulkier than ordinary transformers, and, hence, occupy larger space.
- In a two diode full wave rectifier, only half of the secondary voltage is made use of at a time although it works in both +ve and -ve half cycles.

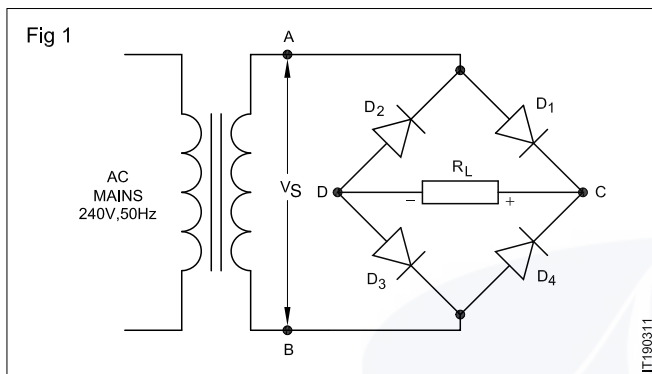
Bridge Rectifiers

Objectives: At the end of this lesson you shall be able to

- describe the working of a bridge rectifier
- state the output DC level in a bridge rectifier
- state the ripple frequency in a bridge rectifier
- state the value of PIV in a bridge rectifier
- compare bridge rectifiers with half-wave and full-wave rectifiers.

THE BRIDGE RECTIFIER

The disadvantages of a full wave rectifier using two diodes and centre-tap transformer can be overcome by a modified full-wave rectifier as shown in Fig 1. In Fig 1, since the diodes are connected in the form of a bridge, this rectifier circuit is commonly known as a *Bridge rectifier*.

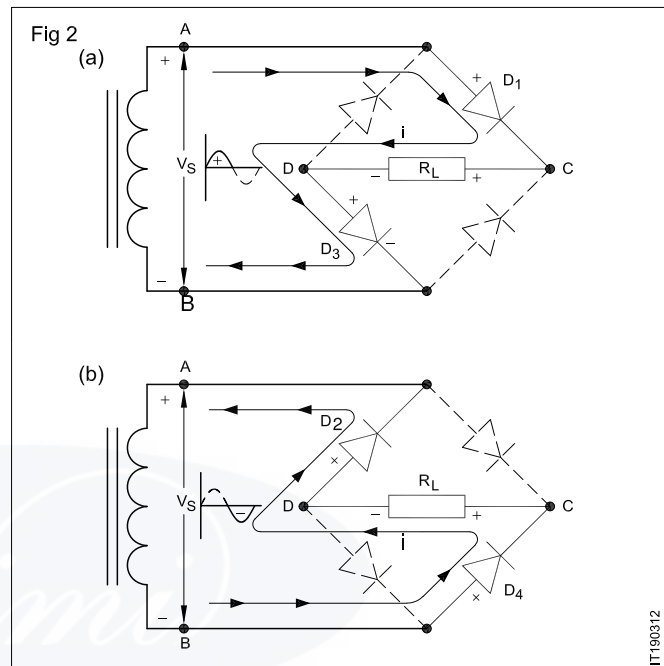


As can be seen in Fig 1, the bridge rectifier does not need a centre-tapped transformer. Also, all the secondary voltage is used for rectification at any given time.

The operation of a bridge rectifier can be summarized in the following steps;

- When end A of the transformer secondary is +ve, as shown in Fig 2a, diodes D_1 and D_3 are forward biased whereas, D_2 and D_4 are reverse biased, and, hence, D_2 and D_4 do not come in the circuit.
- Current flows from the transformer (end A) $\rightarrow D_1 \rightarrow R_L \rightarrow D_3 \rightarrow$ back to the transformer (end B). From the direction of the current flow point C is the positive terminal of the DC output across R_L .
- During the other half cycle of the input (–ve half cycle), end B of the transformer becomes +ve as shown in Fig 2b. Diodes D_4 and D_2 are forward biased, whereas D_1 and D_3 are reverse biased.
- Current flows from the transformer (end B) $\rightarrow D_4 \rightarrow R_L \rightarrow D_2 \rightarrow$ back to the transformer (end A). From the direction of the current flow, point C is again the +ve terminal of the DC output across R_L .

Note that, current I is in the same direction through R_L during both +ve and –ve half cycles of the input AC. The result is, a +ve rectified DC voltage appears at the end of R_L connected to the cathodes of D_1 and D_4 .



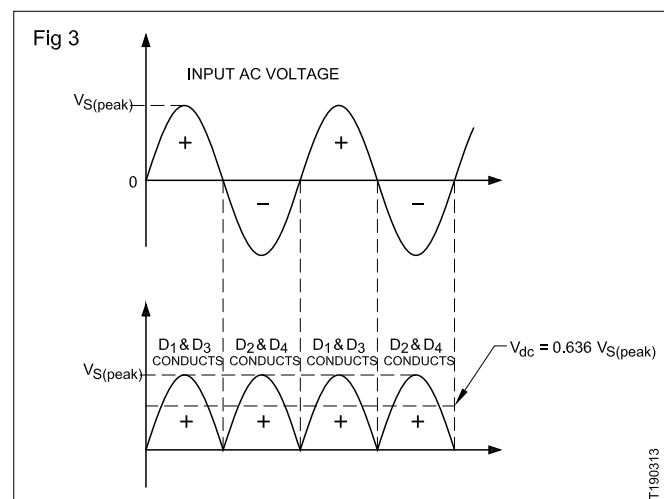
Output DC level in a bridge rectifier

Fig 3 shows the input AC and the output pulsating DC waveform of a bridge rectifier.

This wave-form is similar to that of the full wave rectifier using a centre-tap transformer. Hence, the average DC value of the output is,

$$V_{dc} = 0.636 V_{S(\text{peak})}$$

$$\text{or } V_{dc} = 0.9 V_{S(\text{rms})}$$



where, $V_{S(rms)}$ is the full secondary AC rms voltage.

NOTE: In a two-diode fullwave rectifier $V_{S(rms)}$ refers to only half of the total secondary voltage whereas in a bridge rectifier $V_{S(rms)}$ refers to full secondary voltage.

Example: In Fig 2, if the transformer secondary voltage $V_{S(rms)}$ is 24 volts, the rectified DC voltage V_{dc} across the load R_L will be,

From equation 2, V_{dc} for a bridge rectifier is given by,

$$V_{dc} = 0.9 V_{S(rms)}$$

In the given example, $V_{S(rms)} = 24$ volts

Therefore, $V_{dc} = 0.9 \times 24 = 21.6$ volts

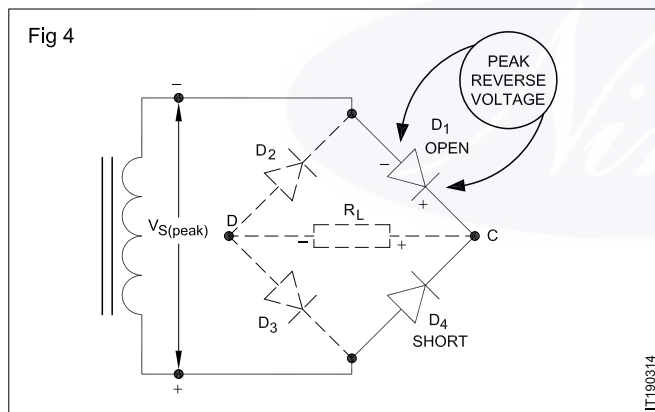
NOTE: Using the same transformer, a two-diode fullwave rectifier would have given only 10.8 volts which is half of that of a bridge rectifier output.

Ripple frequency - Bridge rectifier

The pulsating DC output of a bridge is similar to the two diode fullwave. Hence, as in a two diode fullwave rectifier, the output ripple frequency of the bridge rectifier is also twice the input AC frequency.

Peak inverse voltage - Bridge rectifier

Fig 4 shows a bridge rectifier at the instant the secondary voltage has reached its maximum value.



Diode D_4 is ideally short (as it is conducting) and D_1 is ideally open. Summing the voltages around the outside loop and applying Kirchhoff's law,

$$V_{S(peak)} - \text{PIV across } D_1 + 0 = 0$$

$$\text{or } \text{PIV across } D_1 = V_{S(peak)}$$

Therefore, the peak inverse voltage across D_4 is equal to the peak secondary voltage $V_{S(peak)}$.

In a similar way, the peak inverse voltage across each diode will be equal to the peak secondary voltage $V_{S(peak)}$ of the transformer secondary. Hence the PIV ratings of the diodes used should be greater than $V_{S(peak)}$.

Example: In Fig 4, if the transformer secondary voltage $V_{S(rms)}$ is 24 volts, find the minimum PIV of diodes used.

In a bridge rectifier PIV across the diodes is same and is equal to $V_{S(peak)}$.

Therefore, in the given example,

$$\text{PIV} = V_{S(peak)} = \frac{V_{S(rms)}}{0.707} = \frac{24}{0.707} = 34 \text{ volts.}$$

Current rating of diodes in bridge rectifiers

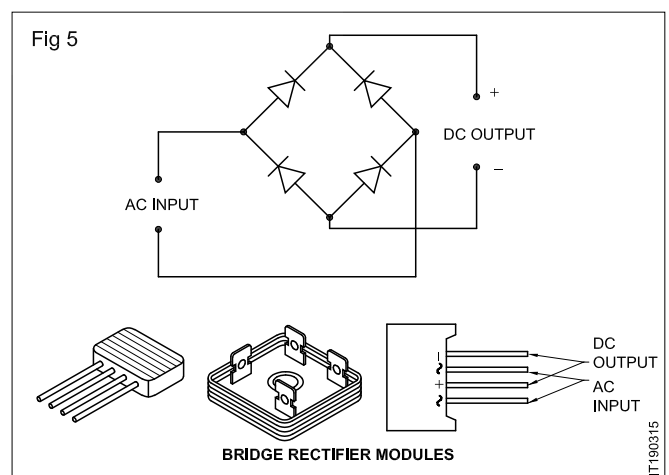
As in the case of a two diode fullwave rectifier even in a bridge rectifier shown in Fig 2, diode pairs D_1, D_3 and D_2, D_4 carry half the total load current I . This is because each diode pair is conducting only during one half of the AC input cycle.

The only disadvantage of bridge rectifiers, if it is treated as a disadvantage, is that, this circuit uses four diodes for fullwave rectification instead of two as in two-diode fullwave rectifier. But this disadvantage is compensated by the simple transformer requirement of the bridge rectifier and higher DC output level. Hence, bridge rectifiers are the most popular AC to DC rectifiers for most applications.

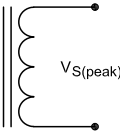
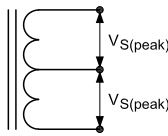
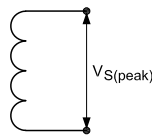
Encapsulated bridge rectifier modules

From the comparative table given at the end of the lesson, it is evident that even though bridge rectifiers need four diodes, it has several advantages over the other two types of rectifiers. Hence, bridge rectifier power supplies are very often used for most applications. Because of its popularity, manufacturers are packing bridge rectifiers as a module for easy use in application. For instance, the B 40 bridge rectifier module is a commercially available bridge rectifier assembly as shown in Fig 5.

The bridge rectifier module shown in Fig 5 consists of four hermetically sealed diodes, interconnected and encapsulated in plastic to provide a single rugged package. It has two input pins for feeding the AC voltage from the transformer secondary and two output pins across which the rectified DC voltage is available.



A comparison of **half-wave**, **fullwave** and **bridge rectifier** is given below in a tabular form;

	Half wave	Full wave	Bridge
• Number of diodes required	1	2	4
• Transformers Peak output voltage			
• DC output voltage in terms of $V_{S(peak)}$	$0.318 V_{S(Peak)}$	$0.636 V_{S(Peak)}$	$0.636 V_{S(Peak)}$
• DC output voltage in terms of $V_{S(rms)}$	$0.45 V_{S(rms)}$	$0.9 V_{S(rms)}$	$0.9 V_{S(rms)}$
• Diode current rating	$I_{L(max)}$	$0.5 I_{L(max)}$	$0.5 I_{L(max)}$
• Peak inverse voltage	$V_{S(peak)}$	$2V_{S(peak)}$	$V_{S(peak)}$
• Ripple frequency	f_{input}	$2f_{input}$	$2f_{input}$

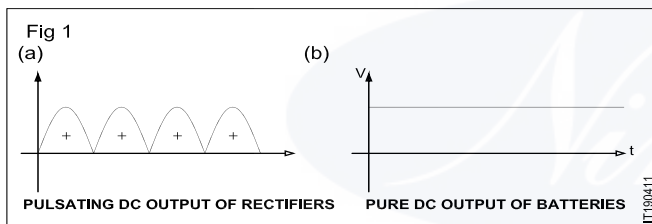
Filters & Rectifiers

Objectives: At the end of this lesson you shall be able to

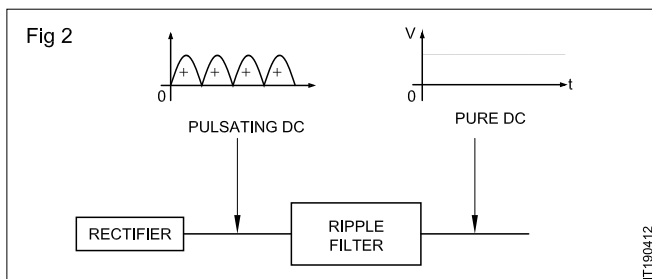
- state the need for filtering pulsating DC
- describe how a capacitor filters the pulsating DC
- state the value of ripple after filtering, using a capacitor input filter in the three types of rectifiers
- state the value of filtered DC output in terms of rms and peak value of input AC in the three types of rectifiers
- design a simple capacitor filter using 10% ripple rule for a full wave rectifier
- state the changed PIV ratings of diodes with capacitor filter
- state the need of surge limiting resistors
- state the advantages and disadvantages of PI filters
- state how inductors can be used to smoothen pulsating DC
- state the output level and the ripple level of a choke input filter
- list popular types of LEDs
- state the advantages of LEDs
- explain why LEDs cannot be confidently tested using ohmmeter
- state how to protect LEDs from high reverse voltage.

RIPPLE FILTERS

The output of rectifiers is a pulsating DC voltage as in Fig 1a and not a pure DC voltage like the output of cells or a battery as shown in Fig 1b.



Pulsating DC voltages cannot be used in most of the electronic circuits like radios, tape recorders etc. These circuits require pure DC voltage similar to that of the output of a battery as shown in Fig 1b. Hence, it is required to remove or at least reduce the pulsation in the output of the rectifier circuits. The circuits used to filter off or reduce the pulsation in the DC output of rectifiers are known as smoothing circuits or more popularly as *ripple filters* as shown in Fig 2.



CAPACITOR INPUT FILTERS

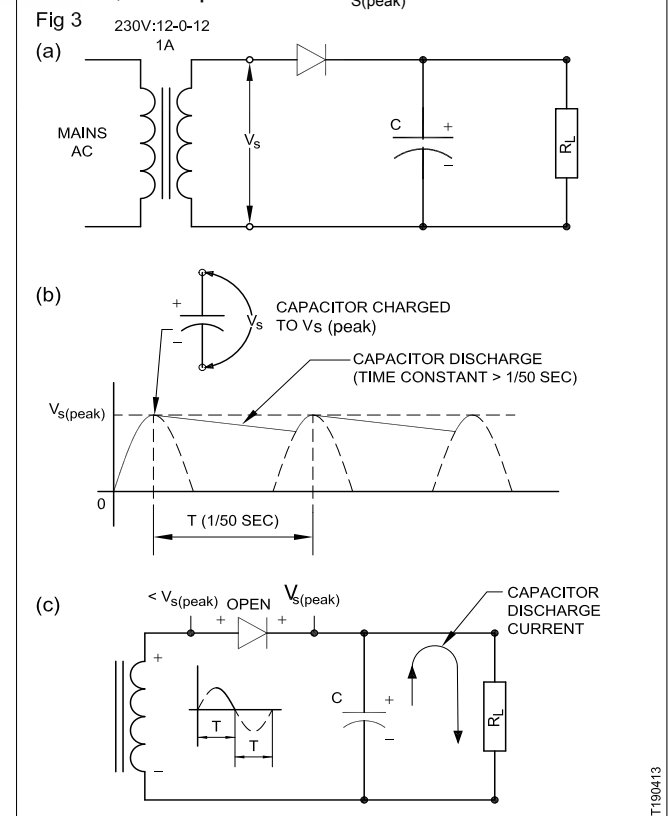
The most important component used in any ripple filter circuit is the capacitor. Recall capacitors store electrical energy and release it when required. This property of a capacitor is made use of in smoothing the pulsating output of rectifiers.

Filtering output of Half-wave rectifiers

Fig 3a shows a capacitor input filter connected at the output of a half-wave rectifier.

During the first quarter cycle of source voltage V_s , the diode is forward-biased and ideally behaves like a closed switch. Since the output of diode is connected directly across the capacitor, the capacitor charges to the peak voltage $V_{s(\text{peak})}$ as shown in Fig 3b.

As shown in Fig 3b, when the input V_s just passes the positive peak, the diode stops conducting. This is because, the capacitor has $V_{s(\text{peak})}$ volts across it with the



polarity shown in Fig 3c. When V_s passes its +ve peak, the voltage at the anode of diode is slightly less than $V_{S(\text{peak})}$. Hence, the diode is reverse biased and behaves like an open switch.

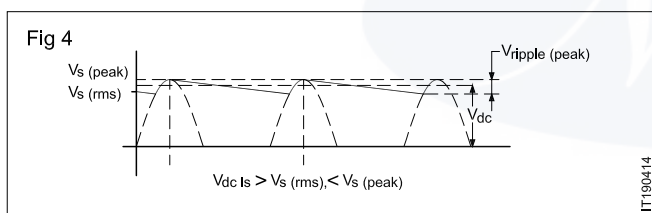
With the diode open, the capacitor discharges through the load resistance R_L . By deliberate design, the discharging time constant (product of R_L and C) is made much greater than the period T of the input signal. Because of this, the capacitor will discharge only a small portion of its stored charge during the OFF time of the diode, as shown in Fig 3b.

By the time the capacitor loses a small portion of its charge, the source voltage again reaches its peak, the diode again conducts and recharges the capacitor to the peak voltage $V_{S(\text{peak})}$. This process of charging during a brief period of positive peak and discharging during the falling period of the positive peak continues as shown in Fig 3b.

From Fig 3b, it can be seen that the voltage across the load resistor is almost a pure DC voltage. The only variation from the pure DC voltage is the small ripple caused by charging and discharging the capacitor.

As the capacitor charges to the peak value of the input $V_{S(\text{peak})}$ and discharges only a small portion of it, the DC output of the half-wave rectifier, instead of $0.45 V_{S(\text{rms})}$ (as discussed in lesson 1.3.02) is only slightly less than $V_{S(\text{peak})}$ as shown in Fig 4.

From Fig 4, the exact value of DC voltage and ripple voltage can be calculated as given below;



$$V_{dc(HW)} = V_{S(\text{peak})} - \frac{V_{rip(p-p)}}{2} \quad \dots\dots[1]$$

Example 1: If $V_{S(\text{peak})} = 34 \text{ V}$ and $V_{rip(p-p)} = 3.4 \text{ V}$,

$$\text{then, } V_{dc} = 34\text{V} - \frac{3.4}{2} = 32.3 \text{ Volts}$$

Refer example in lesson 1.3.02. Without the capacitor the same halfwave rectifier gave only 10.8V DC. Adding just one capacitor the output has increased 3 fold for the same halfwave rectifier.

While designing a filter circuit, the following methods can be used to calculate theoretically the ripple voltage in the output of the filter circuit;

METHOD 1

Knowing the required load current I_L , for a given value of frequency f and capacitance C , the peak-to-peak ripple voltage can be found using the formula,

$$V_{rip(p-p)} = \frac{I_L}{fC} \quad \dots\dots[2]$$

where,

$V_{r(p-p)}$ = peak-to-peak ripple voltage, in volts

I_L = required DC load current, in Amps

f_r = ripple frequency, in H_z

C = capacitance in Farads.

Fixing the permissible $V_{r(p-p)}$ and knowing f and I_L the required value of C can also be found using this formula.

METHOD 2

Another method of expressing the ripple in the output DC is by *ripple factor* r , defined as,

$$\text{Ripple factor, } r = \frac{V_{r(\text{rms})}}{V_{dc}}$$

where,

r = ripple factor (dimensionless)

$V_{r(\text{rms})}$ = rms value of ripple voltage

$$= \frac{V_{rip(p-p)}}{2\sqrt{3}}$$

V_{dc} = DC output voltage.

If the $R_L C$ time constant is large enough such that the capacitor is recharged in a short time compared with half the period T of input AC, the theoretical value of the ripple factor can be calculated using the formula,

$$r = \frac{1}{2\sqrt{3} f R_L C} \quad (\text{for halfwave}) \quad \dots\dots[3]$$

where,

r = the theoretical ripple factor for a half-wave

C = capacitance in μF

R_L = load resistor in ohms

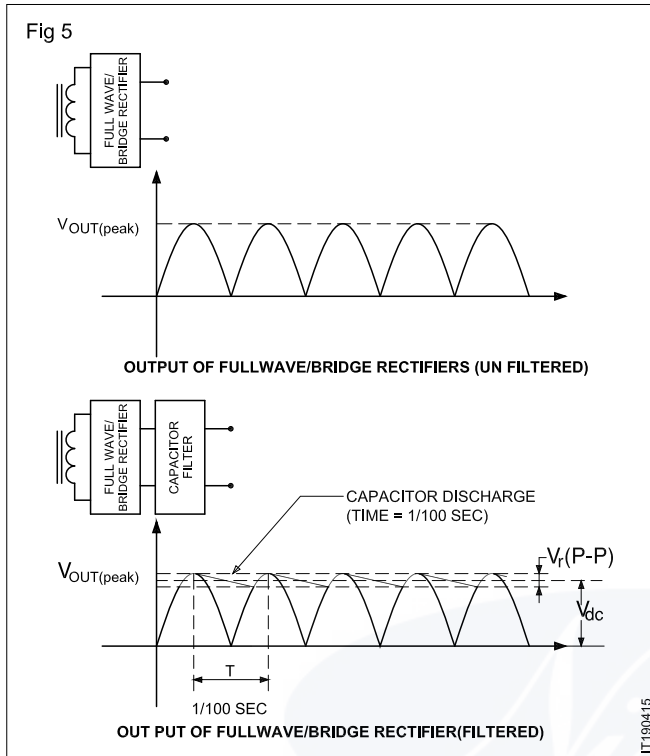
f = mains supply frequency in hertz.

From equation 3 above, it is clear that one way to reduce the ripple is by increasing the discharging time constant, $R_L C$. In other words, increasing the value of C or R_L the time constant becomes larger, and, larger the time constant, purer is the DC output of the filter.

These formulae are used many times while troubleshooting, analyzing, or designing a capacitor-input filter. Hence, it is worthwhile memorizing these key formulae.

Filtering output of Full wave and Bridge rectifiers

From equation 2, it is clear that another way to reduce the ripple is to increase the ripple frequency f_r . In a fullwave rectifier or bridge rectifier, the ripple frequency is double the supply frequency as shown in Fig 5. As a result of higher ripple frequency, when compared with halfwave rectifier (Fig 4), the capacitor gets charged twice as often and has only half the discharge time as shown in Fig 5.



Therefore, the ripple in the filtered output is smaller and the DC output voltage approaches more closely the peak voltage $V_{S(peak)}$.

The formulae used for calculating the ripple in the output or the ripple factor is the same as discussed for half-wave rectifiers.

Example: In a bridge rectifier with a capacitance input filter, suppose the DC load current drawn is approximately 10 mA and the filter capacitor is 470 μ F, for a line frequency of 50 Hz, the peak-to-peak ripple in the output will be,

$$\text{from equation (2), } V_{rip(p-p)} = \frac{I_L}{f_r C}$$

Therefore,

$$V_{rip(p-p)} = \frac{10\text{mA}}{100\text{Hz} \times 470 \mu\text{F}} = 0.213 \text{ V}$$

For the same conditions in a half-wave rectifier, the ripple will be 0.426 V, which is twice as much because the ripple frequency is only 50 Hz (same as AC input frequency).

Ripple filter design guidelines

While designing a capacitor-input filter, you need to choose a capacitor that is large enough to keep the ripple small. How small is the ripple depends on how large is the

capacitor. As capacitor becomes larger, it costs more and hence, more expensive will be the filter circuit.

As a compromise between small ripple and large capacitance, most designs use what is called the *10 percent rule*. This says, select a capacitor that will hold the peak-to-peak ripple at approximately 10 percent of the peak voltage $V_{S(peak)}$.

This level of ripple may look a bit large, but you will see in further units that additional filtering is usually done with electronic circuits called voltage regulators.

For example, if the peak voltage $V_{S(peak)}$ is 34 V, then, select a capacitor that makes the peak-to-peak ripple 10% of $V_{S(peak)}$. This means $V_{r(p-p)}$ should be around 3.4V. (Refer Example 1).

For a bridge rectifier, for a DC load current of 10mA, the value of C using 10% rule should be,

$$V_{r(p-p)} = \frac{I_L}{f_r C}$$

$$\text{or } C = \frac{I_L}{f_r V_{r(p-p)}} = \frac{10\text{mA}}{100(3.4)} = 29.4 \mu\text{F minimum.}$$

As a standard value, use 100 μ F. Recalculating $V_{r(p-p)}$,

$$V_{r(p-p)} = \frac{I_L}{f_r C} = \frac{10\text{mA}}{100 \times 100 \mu\text{F}} = 1 \text{ Volt.}$$

With the 10 percent design rule, the DC voltage is more than 95 percent of the peak voltage which is reasonably a good design.

Calculate value of C if the required load current is 1 Amps.

PIV rating of diodes with capacitance-input filters

Recall that the PIV across the diode in a half-wave rectifier without a filter capacitor is equal to $V_{S(peak)}$. This PIV rating changes when a filter capacitor is connected across the output of the half-wave rectifier.

Fig 6a shows a half-wave rectifier driving a capacitor-input filter. At the negative peak of the secondary voltage, the diode is reverse-biased and a peak inverse voltage appears across it. Summing up the voltages around the loop gives,

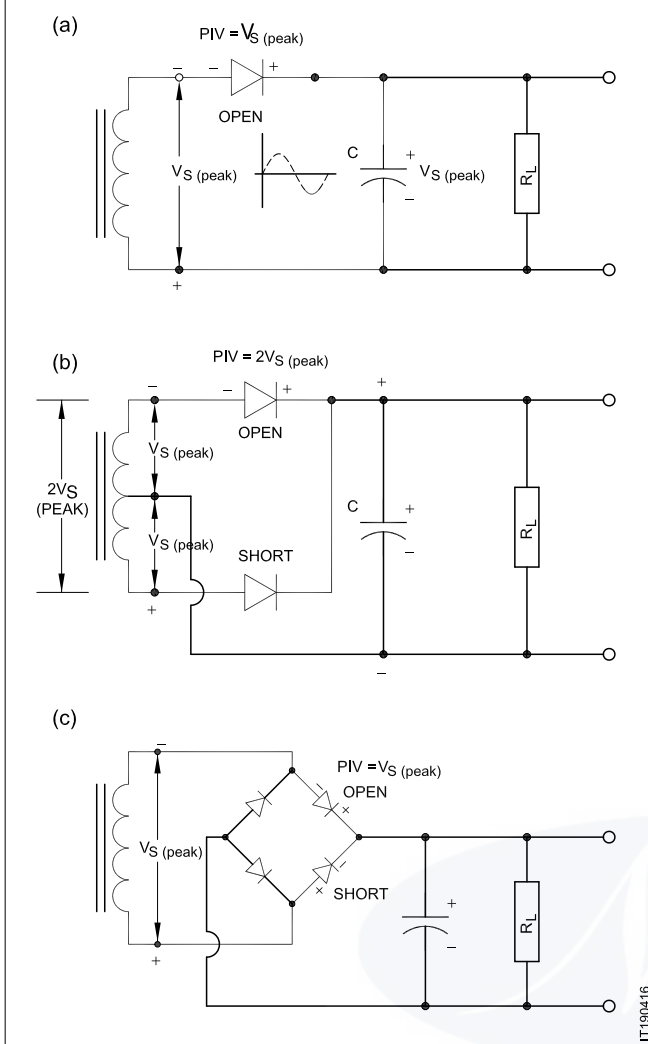
$$\text{PIV across diode} - V_{S(peak)} - V_{S(peak)} \text{ across } C = 0$$

or,

$$\text{PIV across diode} = 2V_{S(peak)}$$

From this it can be seen that, with capacitance-input filter across the diode, the PIV rating of the diode in a half-wave rectifier must be two times the peak secondary voltage $V_{S(peak)}$ or higher.

Fig 6



In the case of a fullwave rectifier with filter shown in Fig 6b, summing voltages around the left loop gives,

$$PIV - 2 V_{S(peak)} + 0 = 0$$

or,

$$PIV = 2 V_{S(peak)}$$

This means that the PIV rating of the diodes must be greater than the peak total secondary voltage ($V_{S(peak)} + V_{S(peak)}$). Note that this is same as in the fullwave rectifier without a filter capacitor.

In the case of a bridge rectifier circuit shown in Fig 6c, summing voltages of the outside loop gives,

$$PIV + 0 - V_{S(peak)} = 0$$

or,

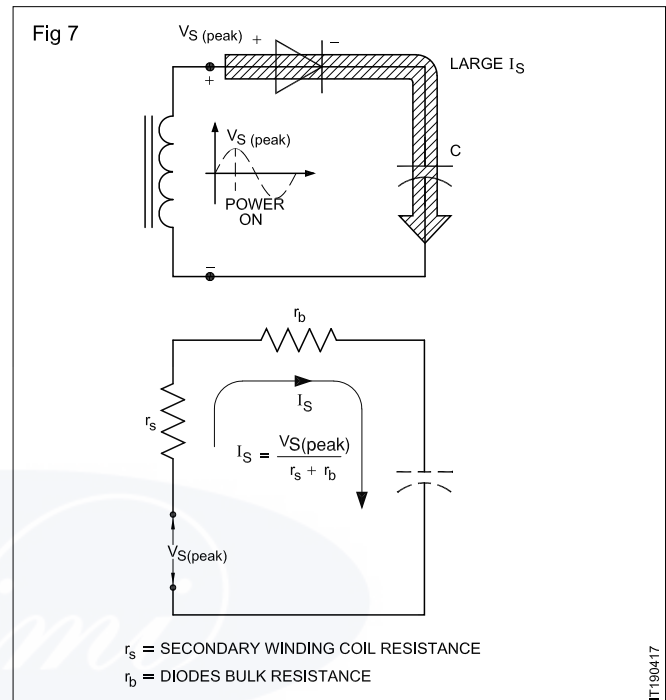
$$PIV = V_{S(peak)}$$

So the PIV rating of each diode in the bridge must be greater than the peak secondary voltage $V_{S(peak)}$. This is same as in a bridge rectifier without a filter capacitor.

Surge current in rectifiers with filter capacitors

In rectifiers with filter capacitors, before the input power is switched ON, the filter capacitor is uncharged. The instant the circuit is switched ON, the capacitor behaves as a short circuit. This results in a large initial charging current through the diode. This initial, sudden rush of current is called the **surge current**.

The worst case of a surge current flowing is when, the circuit is switched ON, at the very instant the secondary voltage is $V_{S(peak)}$, as shown in Fig 7. The only resistance to this rushing surge current are,



- 1 The transformer secondary winding resistance, r_s
- 2 The bulk resistance r_b of the diode(s).

Generally, the secondary winding resistance will be less than 1 W. The bulk resistances of diodes are given in data sheets. In general purpose diodes, r_b will be of the order of 0.2 W.

For example, if $V_{S(peak)}$ is 34 V AC and the sum of the secondary winding resistance and the diode bulk resistance is 1W + 0.2 W, then the surge current I_s would be,

$$I_s = \frac{V_{S(peak)}}{1 \Omega + 0.2 \Omega} = \frac{34}{1.2 \Omega} = 28.3 \text{ amps.}$$

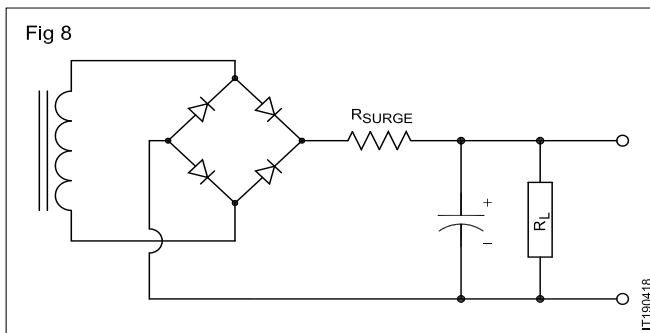
It is important to note that this large current is not a continuous current. This current decreases exponentially as the capacitor gets charged. However, if the capacitor is of an extremely large value, the surge current will remain at a high level for quite some time, and, hence, may damage the diode.

From the above, two important points are to be noted,

- The diode used must have the capability to withstand a large surge current of a short duration, generally one time period(T) of input cycle. This data is given in the diode data books.
- The value of the filter capacitor should not be excessively large as this will increase the value and duration of surge current.

Use of surge limiting resistor

In case the diode selected for a rectifier has a surge current rating less than the estimated surge current, a small resistor is included in the surge current path as shown in Fig 8. This resistor is known as surge limiting resistor or simply surge resistor.



This surge resistor limits the surge current because, this resistance gets added to the winding and diode bulk resistances. Using a surge resistor of a suitable value, the surge current can be restricted to any desired value.

As an example, in Fig 8, if $V_{S(peak)}$ is 34V and if a surge resistor of 4W is selected, then the surge current will be,

$$I_s = \frac{V_{S(peak)}}{\text{winding resistance} + \text{diode bulk resistance} + \text{surge resistor}}$$

$$= \frac{34}{1\Omega + 0.2\Omega + 4\Omega} = 6.5 \text{ Amps}$$

Note that the surge current has reduced from 28.3 amps to 6.5 amps by the use of a surge resistor.

When a surge resistor is used for calculating I_s , you may neglect the winding and bulk resistances, as their values are comparatively smaller.

One disadvantage in using a surge resistor is that, since it comes in series with the output, the DC voltage across the load decreases by an amount equal to the I_R drop across the surge resistor. Hence, surge resistors should not be used indiscriminately.

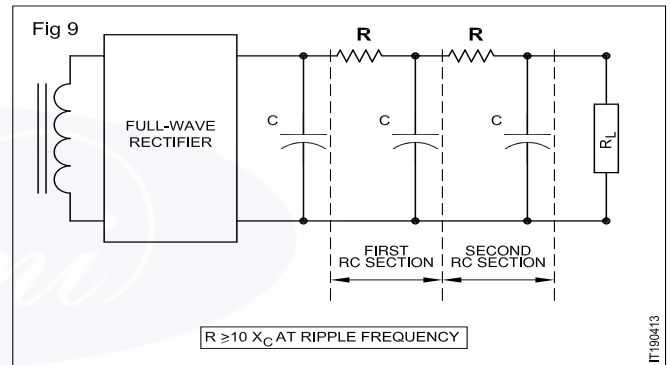
Filter capacitors of values less than 1000 μF usually do not produce a surge current for a long period to damage the diodes. Hence, for values of C less than 1000 μF use of surge resistors are not essential.

Appendix D summarizes the key information discussed so far. This table is useful for designing and troubleshooting power supplies.

RC filters

Recall that with the 10 % ripple rule we get a peak-to-peak ripple of around 10 percent of the DC load voltage. To further reduce the ripple at the output, filters as shown in Fig 9 may be connected between the filter capacitor and the load. These are known as RC p filters. These filters reduce the ripple in the output to less than 1 percent making the output a pure DC similar to the output of a battery.

In Fig 9, the two RC filter sections are connected between the input capacitor C and the load resistor R_L . While designing these filter sections, the value of R is made much greater than X_C at the ripple frequency. Therefore, a major portion of the ripple voltage gets dropped across R instead of across the load resistor R_L . Typically, the value of R must be at least 10 times greater than X_C at the ripple frequency. This results in each RC section attenuating (reducing) the ripple by a factor of 10. Thus more the number of sections of RC, less is the ripple across R_L and purer is the DC output voltage across R_L .



The main disadvantage of the RC filter sections is the loss of DC voltage across each R . Hence, RC filter circuits are suitable only for small load currents or for large load resistance.

Inductor input filters

An alternative to using a capacitor in parallel with the load resistor, is to use an inductor in series with the load as shown in Fig 10a. Since an inductor opposes any change in current through it, the inductor coil will cause a smoothing effect in the output as shown in Fig 10b. Here the inductors are called *Chokes* as they choke the ripple.

If the coil resistance of the inductor is neglected (being very small), the DC output will be,

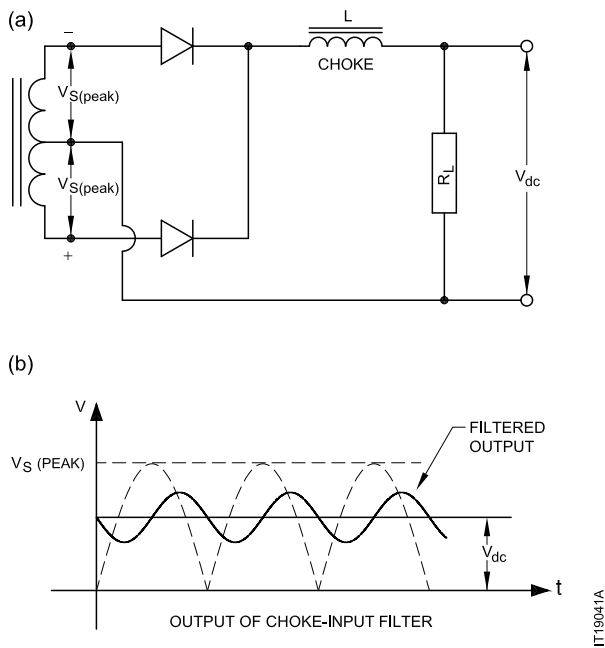
$$V_{dc} = 0.636 V_{S(peak)} \text{ as in the case of capacitance filter.}$$

If the resistance of the coil is considered, the DC output will be slightly less due to the I_R drop across the coil.

The ripple factor r in a full wave rectifier with inductor filter is given by,

$$r = \frac{R_L}{1618L} \dots[4]$$

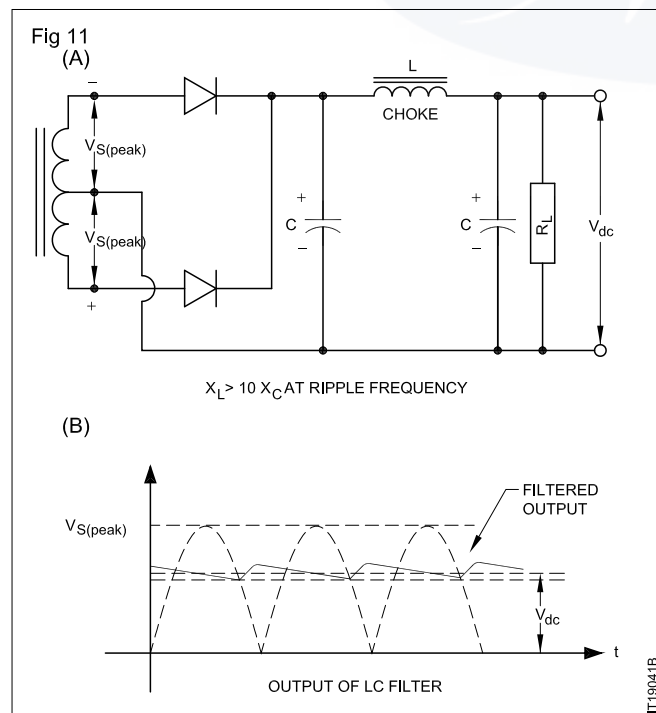
Fig 10



LC filters

For large load current requirements, instead of RC filters, LC filters as shown in Fig 11 are preferred. The reason is, a large portion of the ripple voltage is dropped across the inductor with minimum DC voltage drop because inductors have only a small winding resistance.

In Fig 11, X_L is made much greater than X_C at the ripple frequency. Hence, the ripple across the load is reduced to extremely low levels.



Neglecting the choke's resistance, the DC output voltage is given by,

$V_{dc} = 0.636 V_{S(peak)}$ as in the case of a capacitor input filter.

The ripple factor, r is given by,

$$r = \frac{0.7}{LC} \quad \dots[5]$$

From the above equation for r , unlike in case of capacitor input filter, it can be seen that in a LC filter circuit the ripple factor is not determined by the value of the load resistor. Hence, with a capacitor and inductor in the filter circuit, the ripple is independent of the load current.

These days, LC filters are becoming obsolete because of the size and cost of the inductors. LC filters are being replaced by integrated circuit(IC) voltage regulators, active filter circuits which reduce the ripple and keep the DC voltage constant. IC regulators and active filters are discussed in detail in further lessons.

LIGHT EMITTING DIODES

In recent years, the use of filament lamps/bulbs which consume quite an amount of power, has less life and delicate to handle are becoming obsolete as output indicators of electric systems. With the advancement in the field of optical electronics several devices have been developed as a substitute for filament lamps. One of the most common and popular of these new devices is the **Light Emitting Diode** abbreviated as **LED**. These LEDs are now used as indicators in almost all electrical and electronic circuits and equipments.

The advantages of LEDs over incandescent bulbs are listed below;

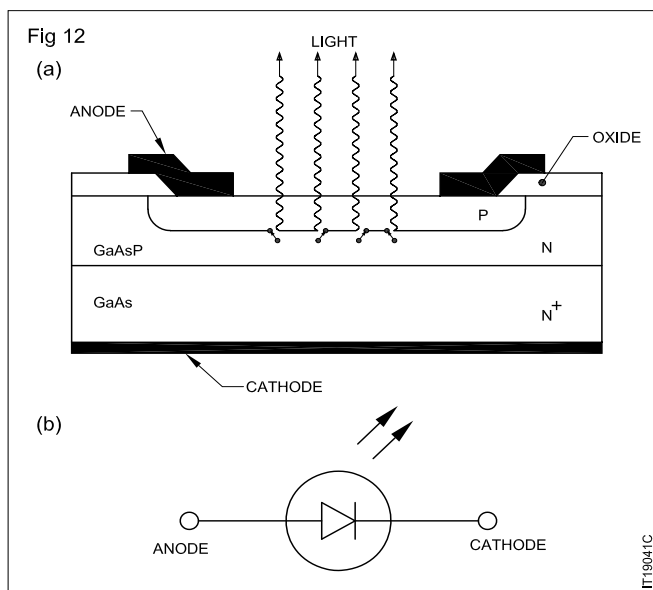
- 1 LEDs have no filaments to heat and so require less current to glow.
- 2 LEDs require lower voltage level (typically 1.2 to 2.5 V) than the conventional bulbs.
- 3 LEDs last much longer - upto several years.
- 4 Because there is no filament to heat up, LEDs are always cool.
- 5 LEDs can be switched ON and OFF at a much faster rate compared with conventional lamps.

Principle of working of LEDs

LED is nothing but a type of diode. LEDs also possess the unique unidirectional property like general purpose diodes. But, the materials used in making LEDs are different; hence, their characteristics are different too. Therefore, it is very important to note that *although LED is also a type of diode, it cannot and should not be used for the purpose of rectifying ac to dc.*

Recall that a general purpose diode or a rectifier diode conducts when energy is supplied to the electrons ($Si=0.7V$, $Ge=0.3V$) to cross the barrier junction. Each electron, after acquiring the supplied extra energy, crosses the junction and falls into the hole on the P side of the junction while the electron recombines with a hole, the electron gives up the extra energy by it. This extra energy is dissipated in the form of heat and light. In general purpose diodes because the silicon material is not transparent(opaque), the light produced by the electrons does not escape to the outer environment. Hence, it is not visible.

But LEDs are made using semi-transparent materials instead of silicon. Because the material used in making LEDs is semi-transparent, some of the light produced by the electrons escapes to the surface of the diode, and, hence, is visible as shown in Fig 12a.



LEDs are typically doped with gallium arsenic, gallium phosphate or gallium arseno-phosphate. Different dopes cause the LED to emit light of different colours (wavelengths) such as red, yellow, green, amber, or even invisible infrared light.

The schematic symbol of LED is shown in Fig 12b. The arrows are used to indicate that light is radiated from the device.

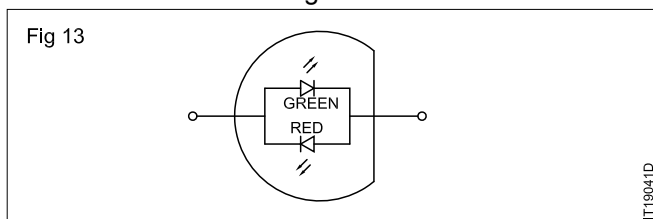
Types of LEDs

Single colour LEDs: Most of the commercially available and commonly used LEDs are single colour LEDs. These LEDs radiate one of the colours such as red, green, yellow or orange. Different coloured LEDs will have different forward voltages as given in the table below:

Colour of LED	Red	Orange	Yellow	Green
Typical Forward voltage drop	1.8V	2V	2.1V	2.2V

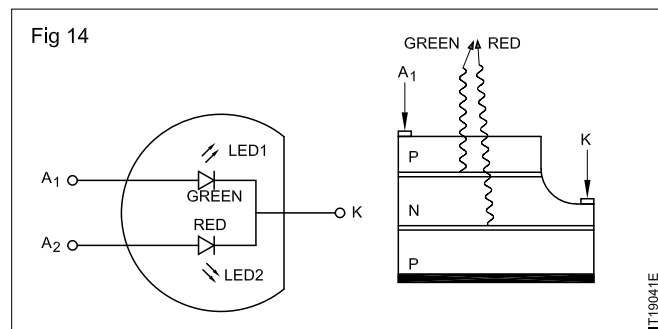
These typical forward voltage drops are at a typical LED forward current $I_f = 20 \text{ mA}$

Two colour LEDs: These LEDs can give two colours. Actually, these are two LEDs put in a single package and connected as shown in Fig 13.



In a two-colour LED, two LEDs are connected in inverse parallel, so that one of the colour is emitted when the LED is biased in one direction and the other colour is emitted when the LED is biased in the other direction. These LEDs are more expensive than the single colour LEDs. These LEDs are useful to indicate +ve, -ve polarities, GO-NOGO indication, null detection etc.

Multicolour LEDs: These are special types of LEDs which can emit more than two colours. These LEDs comprises of a green and a red LED mounted in a three-pin common cathode package as shown in Fig 14.

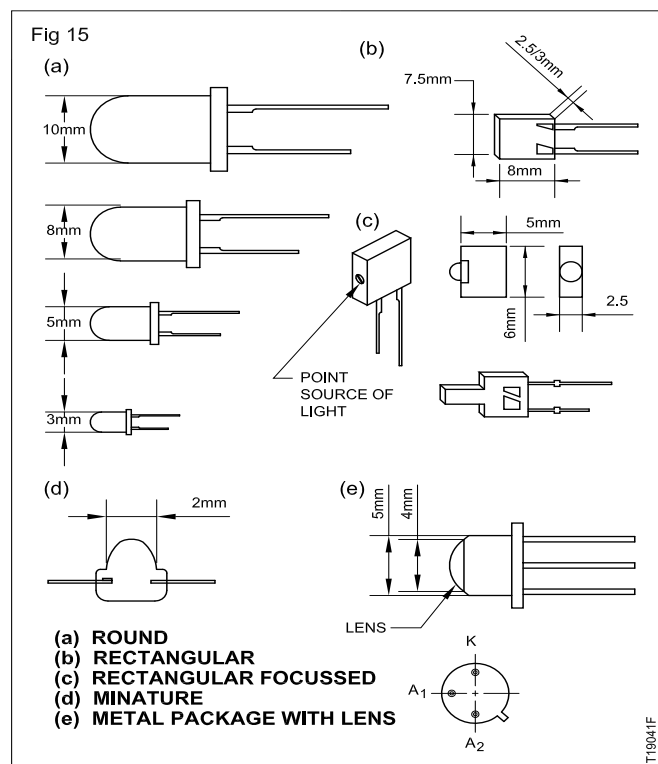


This LED will emit green or red colour by turning ON only one LED at a time. This LED will emit orange or yellow by turning on the two LEDs with different current ratios as shown in the table given below:

Output colour	Red	Orange	Yellow	Green
LED-1 current	0	5mA	10mA	15mA
LED-2 current	15mA	3mA	2mA	0

Sizes and shapes of LEDs

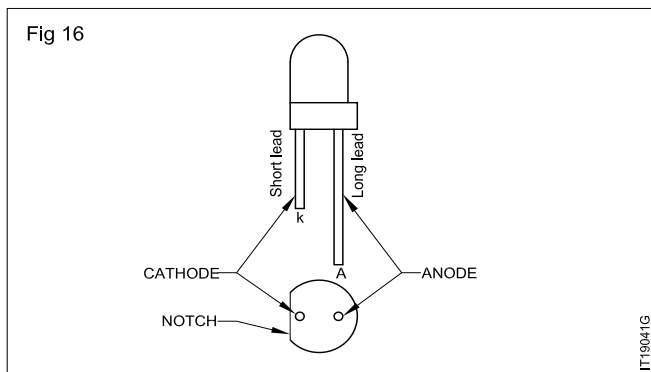
LEDs are available commercially in different shapes and sizes to suit varied commercial applications. Fig 15 shows some of the most popular shapes and sizes of LEDs.



The light output of LED may be guided as point-source or diffused. The point-source LED provides a small point of light while the diffused type has a lens which diffuses the light into a wide angle viewing area.

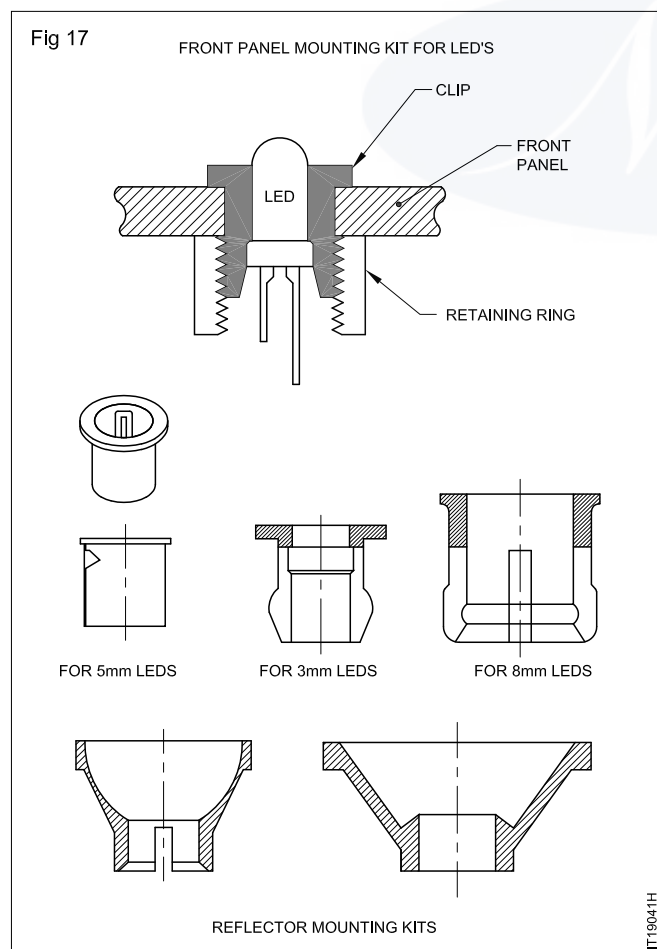
Terminals of LEDs

Since LEDs are basically diodes, they have anode and cathode terminals/leads as in any general purpose diode. Fig 16 shows the methods to identify the terminals of a LED.



Mounting kits for LEDs

Special mounting kits, as shown in Fig 17, are available for fixing the LEDs on to the printed circuit boards and monitoring panels. These kits not only extend the life of the LED by way of protecting it from mechanical stress but also make the output of the LED clearly visible.



Testing LEDs using Ohmmeter

The anode and cathode terminals of a general purpose diode can be checked easily using an ohmmeter. But, in the case of LEDs, unlike general purpose diodes, the forward voltage of LED ranges from 1.5 to 3 volts (in some cases it is higher than 3 V), and a typical forward current ranges from 10 mA to more than 50mA. Because of this large forward voltage and current requirement of the LEDs, it is not always possible to test the LEDs using an ohmmeter. The reason is, most of the portable type ohmmeters/multimeters use internal battery of not more than 3V for the operation of the meter. This voltage would have got reduced with constant use of the meter. Hence, when an LED is tested using an ohmmeter, the glow of the LED may be very dim or the LED may not glow at all depending on the condition of the battery inside the meter. Hence, the condition of an LED cannot be confidently confirmed using a meter. However, since meter testing is the quickest, this can be used while purchasing an LED from the vendor where other equipments may not be available for testing.

Specifications of LEDs

Specifications sheet of a typical LED is given in the table below;

A Typical LED-specification sheet

(For: FairChild, FLV117 Red LED)

Characteristics	Min.	Typical	Max.
Forward current, I_f		20 mA	50 mA
Forward voltage, V_f		1.7V	3V
Reverse voltage, V_R		8V	
Axial luminous intensity	0.8 mcd	2 mcd	
Angle of half intensity		$\pm 20^\circ$	
Peak wave-length		665 nm	

From the specifications of a typical LED given above, the following important points are to be noted;

- The forward voltage drop of the LED is much higher (1.7V to 3V) than that of general purpose diodes.
- The reverse voltage that can be applied to the LED is much lower than in general purpose diodes.

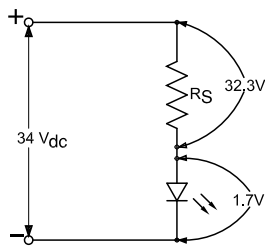
The above two important points confirm that, LEDs do not have the same characteristics as general purpose diodes.

In the typical LED specification, for instance, if 8 V or more is applied across the LED in the reverse biased polarity, the LED will be destroyed.

Example: What value of R_s is required, if a red colour LED is to be used in a circuit with a source of 34V dc.

Using the specifications of the red LED given in the table, it is clear beyond doubt that the LED cannot be connected across 34 volts supply directly (maximum $V_f = 3V$). Hence, as shown in Fig 18, a resistor is to be used in series with the LED which must drop to 32.3 volts if the voltage across LED should be 1.7 V.

Fig 18



IT190411

For the LED to give reasonably good light, the current through the LED has to be 20 mA, as indicated in the specifications sheet. So, the value of R_S must be,

$$R_S = \frac{V}{I} = \frac{32.3 \text{ V}}{0.02 \text{ A}} = 1615 \, \Omega$$

Since the maximum permissible current through the LED is given as 50 mA, it is possible to use a standard 1.6KW resistor. This will make a current of 20.2 mA to flow through the LED which is well within the permitted maximum current rating.



Characteristics & specification of zener diodes

Objectives : At the end of this lesson you shall be able to

- state the need of regulators in power supplies
- state the formula to calculate the % load regulation factor
- list the main differences between rectifier diodes and zener diodes
- list the similarities between rectifier diodes and zener diodes
- name the main application of zener diodes
- list the important specifications of a zener diode
- interpret the specifications of some zener types without referring to data book.

Zener Diode

A Zener diode is a special purpose diode that is designed to operate in the reverse breakdown region of the diode's characteristic curve. Regular diodes will be destroyed if they are used in the reverse breakdown region.

Zener Diode Symbols

The schematic and physical symbols are shown. Note that the symbol is similar but somewhat different than that of a regular diode.

The Zener diode is off when the applied voltage is less than the Zener voltage rating. When the applied voltage is greater than the Zener voltage rating the Zener diode turn on and current flows.

Zener Diode Specifications

Zener (Breakdown Voltage) Voltage - V

V_Z is the most common specification. Zener diode voltages range from 3.3 V to 100 V or more. As seen in the graph, the Zener voltage does increase a small amount as the diode current increases. This reflects the fact that the line in the graph is not perfectly vertical.

Power Dissipation – P₀

This is the rating for the maximum amount of power the Zener diode can dissipate. Values range from 0.3 W to 50 W.

Power Derating Factor - PDF

The maximum power that the Zener can safely dissipate decreases with temperature. The Power Derating Factor specifies this decrease in maximum power rating.

Voltage regulators

Recall that, the DC output voltage level of power supplies such as, full-wave and bridge rectifiers, tend to decrease or increase,

- when the load current increases or decreases
- when the AC input voltage level decreases or increases.

Such variations in the output DC voltage level of power supply is not acceptable for most of the electronic circuits. Hence, it is required to regulate the DC output of power

supplies so as to keep the DC output level constant, inspite of variations in the DC load current or the AC input voltage. Circuits or components used to keep the DC output voltage of a power supply constant are called voltage regulators.

Regulation factor

The ability of a power supply to maintain a constant DC output voltage for variations in the load current is referred to as load regulation. Load regulation of a power supply is generally given as a percentage.

$$\text{Load regulation factor \%} = \frac{V_{NL} - V_{FL}}{V_{NL}} \times 100$$

where,

V_{NL} = DC output at no load or open circuit

and V_{FL} = DC output at rated full load.

It should be noted that lower the percentage of load regulation factor, better is the voltage regulation.

Example: The DC output of a power supply is 12 volts at no-load and 11 volts at full load.

$$\% \text{ Load regulation} = \frac{12 - 11}{12} \times 100 = 8.33\%$$

In practice the load regulation of a good power supply should be less than 0.1 %.

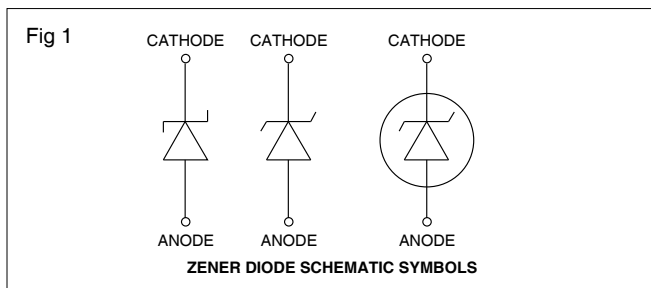
Regulating the DC output voltage for variations in the input AC level is termed as line regulation. This is discussed in further units.

The zener diode

In a power supply one of the simplest ways of regulating the DC output voltage (keeping the output voltage constant) is by using a zener diode. With zener in reverse breakdown condition, the voltage across the zener diode remains constant for a wide range of input and load variations.

Because of this property, zener diodes are also known as voltage regulators or voltage reference diodes. Fig 1 shows the symbol used for zener diodes.

The difference between a rectifier diode and a zener diode are listed below;



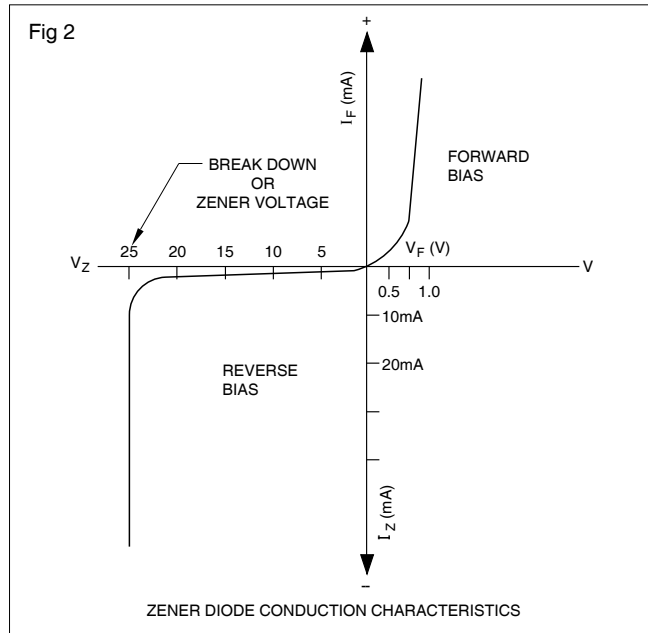
- Compared to normal rectifier diodes, zener diodes are heavily doped.
- Unlike ordinary diodes which do not work in the breakdown region, zener diodes work only in the breakdown region.
- General rectifier diodes are used in forward-biased condition, whereas zeners are always used in reverse-biased condition.
- The reverse breakdown voltage of zener diodes is very much less (3 to 18V) compared to rectifier diodes (minimum 50V).

The similarities of a zener diode with those of general purpose rectifier diodes are listed below;

- Zener diodes are also PN junction diodes, which are also generally made of silicon.
- Zener diodes also have two terminals (anode and cathode).
- In physical appearance, the zener diodes and ordinary diodes look alike.
- Like rectifier diodes, zener diodes are also available with glass, plastic and metal casing.
- The anode and cathode marking technique on the body is same for both zener and rectifier diodes.
- The zener can be tested with an ohmmeter in the same way as in rectifier diodes.
- Zener requires approximately the same voltage for it to be forward-biased into conduction as that of an ordinary diode.

Fig 2 shows the conduction characteristics of a typical zener diode. Because of the nature and heavy doping in a zener, its characteristics are different compared to a rectifier diode.

Note that, the zener diode acts as a rectifier diode when forward biased. It also behaves as a rectifier diode when reverse-biased, till the voltage across it reaches the breakdown voltage. As can be seen from Fig 2, even the reverse or leakage current remains almost negligible and constant despite the increase in the reverse-biased voltage till the break down voltage, also called zener voltage is reached. But, Once the zener breakdown voltage is reached, the diode current begins to increase rapidly and the zener suddenly begins to conduct. In the case of a normal rectifier diode, once the break down voltage is reached the diode gets punctured and starts conducting heavily whereas, in a zener diode, the diode does not get punctured even though it conducts current in the reverse biased condition.



The cause for this reverse conduction is referred to as the avalanche effect. The avalanche effect cause, the electrons to be knocked loose from their bonds in the crystal structure. As more electrons are loosened, they in turn knock others and current builds quickly. This action causes the voltage drop across the zener to remain constant regardless of the zener current. As shown in Fig 2, once the zener voltage is reached, very small voltage changes create much greater current changes. It is this characteristic, which makes the zener useful as a constant voltage source or as a voltage regulator.

Unlike in a rectifier diode, the reverse current through the zener is not destructive. If the current is kept within the specified limits depending upon the wattage rating of the zener, using a suitable series resistance, no harm is done to the zener diode.

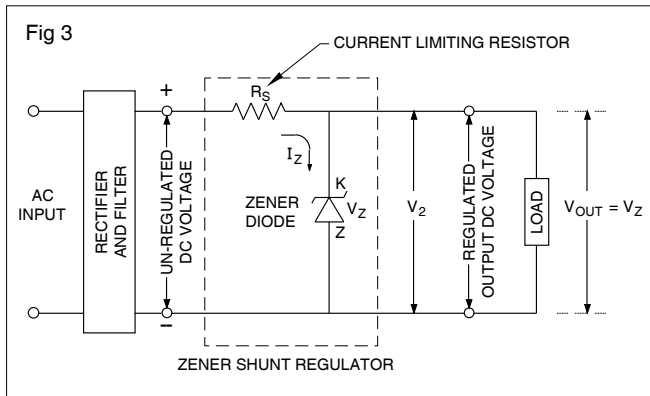
Because the zener diode is designed to operate as a breakdown device, the zener can be brought out of condition easily. A zener is brought out of its zener conduction by lowering the reverse-biased voltage below the zener voltage or by reversing the polarity of the applied voltage.

Application of zener diodes

The most popular use of zener diodes is as voltage regulators in DC power supplies. Fig 3 illustrates a simple zener regulated power supply.

In the circuit at Fig 3, the zener diode is in parallel with the output or load of the power supply. It is very important to note that the zener is connected in the reverse-biased condition. Such a parallel circuit connection is often called a shunt. When used in this way, the zener is said to be a shunt regulator.

In Fig 3, the zener begins to conduct in the reverse-biased condition as the voltage across it reaches the zener voltage V_Z . The voltage across the zener remains constant immaterial of the input DC voltage. Since the load is in parallel with the zener, the voltage across the load V_{OUT} will be same as the voltage across the zener V_Z ($V_{OUT} = V_Z$).



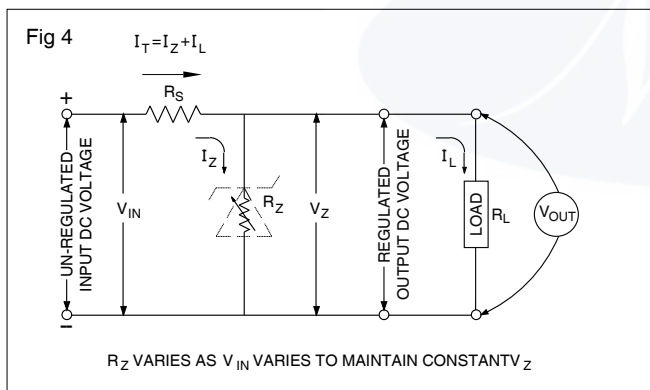
If the input DC voltage to the zener increases, as can be seen from its characteristics in Fig 2, the current I_Z through the zener increases but the voltage across the zener remains the same due to avalanche effect. Because the zener voltage, V_Z does not change, the output voltage V_{OUT} , does not change and so the voltage across the load is constant. Thus, the output is said to be regulated.

Referring to Fig 4, the zener can be looked at as an automatically changing resistance. Total current through the resistance R_S is given by,

$$I_T = I_Z + I_L$$

Thus the voltage across R_S is,

$$V_R = (I_Z + I_L) R_S$$



If the input DC voltage V_{IN} increases, output voltage V_{OUT} tends to increase. In the meantime, the zener conducts more heavily, causing more current (more I_Z) to flow through R_S . Hence, more voltage drop occurs across R_S . This increase in drop across R_S offsets the increase in the output voltage V_{OUT} , thus retaining the voltage across load R_L at its original value. Likewise, if the value of R_L is decreased (increased I_L), current through the zener I_Z decreases, retaining the value of I_T through R_S . This ensures sufficient load current through the load R_L without decrease in the level of V_{OUT} .

Zener specifications

Like in rectifier diodes, the type-code number is marked generally on the body of the zener. From the type-code marked, detailed specifications of the zener can be found referring to any standard diode data manual.

Important zener diode specifications are listed below;

- **Nominal Zener voltage, V_Z :** This is the reverse biased voltage at which the diode begins to conduct in reverse bias.
- **Zener voltage tolerance:** Like the tolerance of a resistor, this indicates the percentage above or below V_Z . For example, $6.3 \text{ V} \pm 5 \text{ percent}$.
- **Maximum zener current, $I_{Z,max}$:** This is the maximum current that the zener can safely withstand while in its reverse-biased conduction (zener) mode.
- **Maximum power dissipation, P_Z :** is the maximum power the zener can dissipate without getting damaged.
- **Impedance (Z_Z):** The impedance of the zener while conducting in zener mode.
- **Maximum operating temperature :** The highest temperature at which the device will operate reliably.

These specifications of zener diodes are given in diode data books. However a limited list of the most commonly used zener diodes is given in Appendix D

The example given below enables you to interpret the specifications of certain types of zener diodes without the need to refer diode data book:

Example 1: The type-code printed on a zener is BZC9V1.

BZC9V1

B	Z	C	9V1
silicon	zener	5% tolerance	9.1V

Example 2: The type-code printed on a zener is 1Z 12.

1Z 12

1	Z		12
Means a semiconductor with one PN junction	zener	No tolerance code means, 10% tolerance	12V

Other popular zener diode type-codes are, 1N750, 1N4000, ZF27, ZP30, DZ12, BZ148, Z6, etc.

Zener diodes - Designing regulators

Objectives: At the end of this lesson you shall be able to

- state the need for minimum current through the zener
- state the worst case conditions to be considered while designing a zener regulator
- calculate the value and wattage of a series resistor
- calculate the required wattage of a zener for a given application.

Designing a simple zener regulator

In the lesson earlier it was discussed that if the voltage across a zener tends to increase or decrease, it results in increase or decrease in the current I_z through the zener. This variation in I_z results in the voltage across the series resistor R_s to increase or decrease, thus keeping the voltage across the zener, and, hence, the voltage across the output/load constant.

From this it is clear that, to make a voltage regulator circuit using zener, a resistor and a zener diode are required. The value of the resistor should be chosen so as to meet the following conditions;

- Under full load condition (i.e., $I_L = \text{max}$), at least the minimum reverse current should flow through the zener such that the zener remains to be in the zener breakdown condition.

The voltage drop across R_s should be such that,

$$V_z + V_{RS} = V_{IN}$$

where, V_z is the zener voltage and also the required output voltage V_{OUT} of the regulator.

For the zener to hold the output voltage constant, the zener must remain in the breakdown region under all conditions.

- Under no load condition, the series resistor R_s must restrict the current through the zener, such that, the power dissipation across the zener is within the specified limit of the device.

The voltage drop across R_s should be such that,

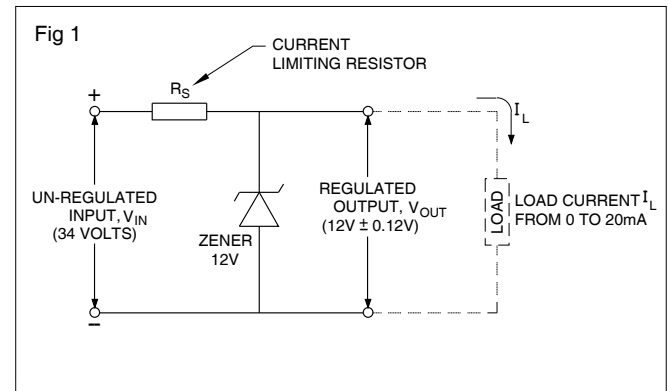
$$V_z + V_{RS} = V_{IN}$$

The design steps for a simple zener regulator circuit is given below through an example:

Example: A zener regulator circuit is needed to supply a constant output voltage of 12V DC $\pm 0.1V$. The load current may vary (depending on load resistance) from 0 to 100mA. The unregulated input to the regulator is 34V DC (maximum).

Design steps:

- 1) Draw a schematic of regulator as shown in Fig 1.
- 2) Choose a zener of $V_z = 12$ volts as the output voltage required is 12 Volts. Choose the zener with a tolerance of $\geq 10\%$, so that the output will be 12 V DC ± 0.12 V.



- 3) From the specifications of the zener, find I_z . Say the I_z of the chosen zener is = 20mA.

- 4) Calculate the current through the zener in the worst conditions as explained below;

One of the worst conditions is, when the input voltage V_{IN} is minimum and, the load current is maximum. For this condition, choose the minimum I_z that should flow through the zener to keep it in reverse-ON condition.

In the example considered, $I_z = 20\text{mA}$.

$$\text{Since, } I_T = I_z + I_{L(\text{max})}$$

For the given example,

$$I_T = 20\text{mA} + 100\text{mA} = 120\text{mA}.$$

The other worst condition is, when maximum current flows through the zener as the load current is zero or minimum and the source voltage is maximum. In the example considered, minimum $I_L = 0\text{mA}$.

When $I_L = 0\text{mA}$, current through the zener will be maximum and is,

$$120\text{mA} - 0\text{mA} = 120\text{mA}.$$

- 5) Calculate the zener wattage.

The zener wattage is determined from the maximum I_z

$$W_z = V_z \times I_z = 12 \times 0.12 = 1.44\text{ Watts}$$

where, V_z is the zener voltage and I_z is the maximum current through zener.

6) Calculate the required value of R_S

$$V_{IN} - V_{OUT} = I_T R_S$$

or,

$$I_T R_S = V_{IN} - V_{OUT}$$

$$= 34V - 12V = 22V$$

Since, $I_T = 120mA$

$$R_S = \frac{22}{120 \times 10^{-3}} = 183\Omega$$

Therefore, $R_S = 183\Omega$

Since 183 is not a standard resistance value, choose the nearest lower value of R_S which is 180Ω upon choosing 180Ω in place of 183Ω , the changed value of I_T will be,

$$I_T = \frac{22V}{R_S} = \frac{22V}{180\Omega} = 122mA$$

The increased I_T of 122mA instead of 120 results in an excess of 2mA through the zener. Therefore add an extra wattage of $12V \times 0.002 = 24mw$ to the already calculated zener wattage of 1.44 watts. Therefore the new wattage of zener should be $1.4 \text{ watts} + 0.024 \text{ W} = 1.464 \approx 1.5 \text{ W}$.

7) Calculating the wattage of R_S

Maximum current through R_S in the worst condition will be 122mA.

Therefore, $W = I_T^2 R_S = (0.122)^2 \cdot 180$

$$= 2.69 \approx 2.7 \text{ watts}$$

From the above calculations, On No-load,

$$I_Z + I_L = I_T = 122mA$$

Since $I_L = 0$, $I_Z = I_T = 122mA$

On Full-load (100mA)

$$I_T = I_Z + I_L$$

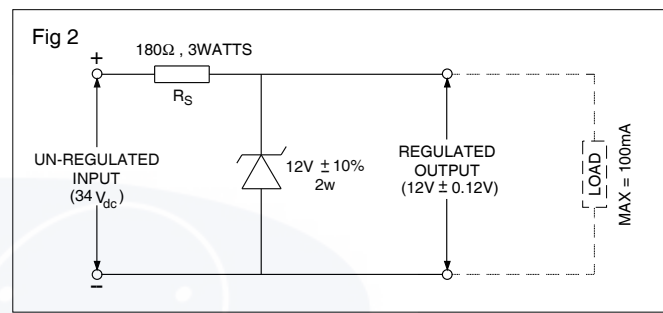
$$I_T = 122mA$$

Therefore, $I_Z = I_T - I_L$

$$= 122mA - 100mA = 22mA$$

I_Z of 22mA is sufficient to keep the zener in reverse ON condition even during full load.

In steps 6 and 7, a good design practice is to overrate the wattage of the resistor and zener, so that the components do not get overheated and fail prematurely. Hence, the resistor should be rated at 3 watts or more and the zener at 2W. The designed simple zener regulator with component values and ratings is shown in Fig 2.



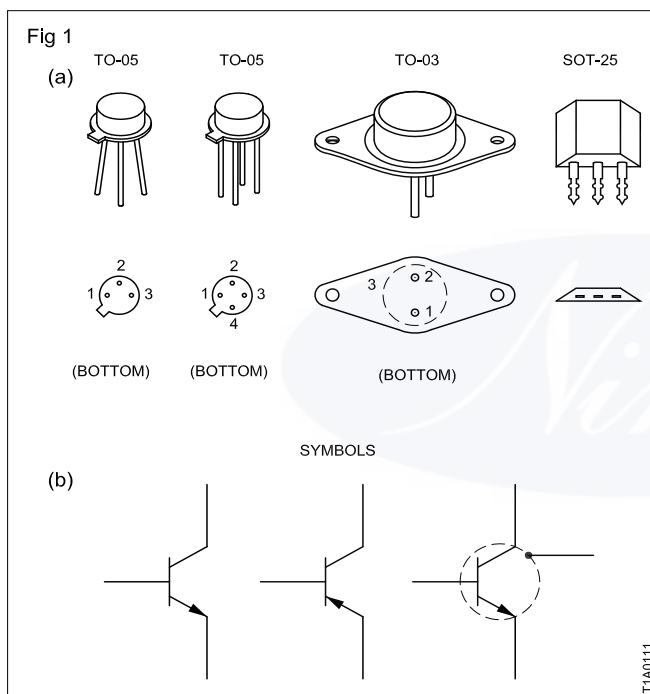
Identify the test different types transistors

Objectives: At the end of this lesson you shall be able to

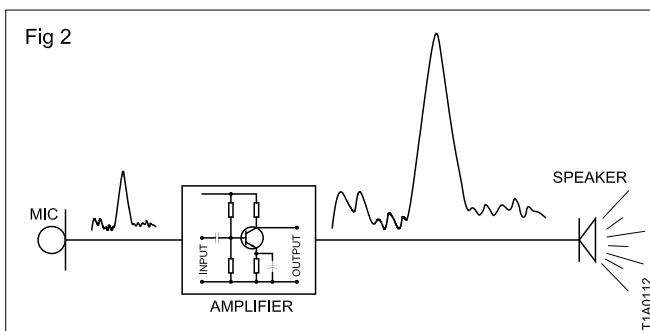
- state the two main uses of transistors
- list the advantages and classifications of transistors over vacuum tubes
- state the use of a transistor data book
- state the names given to the leads of a transistor
- state the functions of the three sections of a transistor
- describe the two tests to be conducted on a transistor before using it.

INTRODUCTION TO TRANSISTORS

Transistors are the semiconductor devices having three or four leads/terminals. Fig 1a shows some typical transistors. Fig 1b shows the symbols used for different types of transistors.

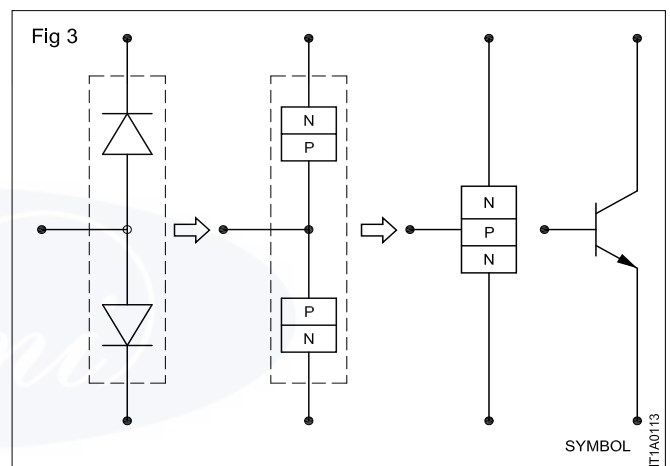


Transistors are mainly used for enlarging or amplifying small electric/electronic signals as shown in Fig 2. The circuit which uses transistors for amplifying is known as a transistor amplifier.

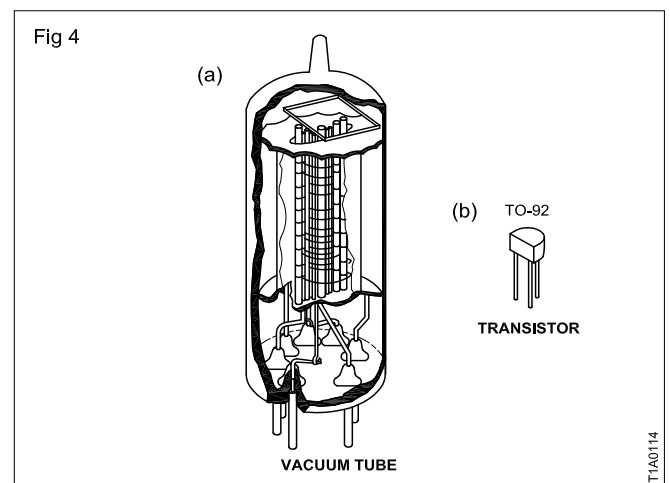


One other important application of transistors is its use as a solid state switch. A solid state switch is nothing but a switch which does not involve any physical ON/OFF contacts for switching.

Transistors can be thought of as two PN junction diodes connected back to back as shown in Fig 3.



Before the transistors were invented (1947), there was what were known as vacuum tubes which were used in amplifiers. A typical vacuum tube is shown in Fig 4a.



Compared with the present day transistors the vacuum tubes were big in size, consumed more power, generated lot of unwanted heat and were fragile. Hence vacuum tubes became obsolete as soon as transistors came to market.

Transistors were invented by Walter H. Brazil and John Barlow of Bell Telephone Laboratories on 23rd Dec. 1947. Compared to vacuum tubes (also known as valves), transistors have several advantages. Some important advantages are listed below;

- Very small in size
- Light in weight
- Minimum or no power loss in the form of heat
- Low operating voltage
- Rugged in construction.

To satisfy the requirements of different applications, several types of transistors in different types of packaging are available. As in diodes, depending upon the characteristics, transistors are given a type number such as BC 107, 2N 6004 etc., The characteristics data corresponding to these type numbers are given in Transistor data books.

CLASSIFICATION OF TRANSISTORS

1 Based on the semiconductor used.

- Germanium transistors
- Silicon transistors

Like in diodes, transistors can be made, using any one of the above two important semiconductors. However, most of the transistors are made using silicon. This is because, silicon transistors work better over a wide temperature range (higher thermal stability) compared to germanium transistors.

Transistor data books give information about the semiconductor used in any particular transistor.

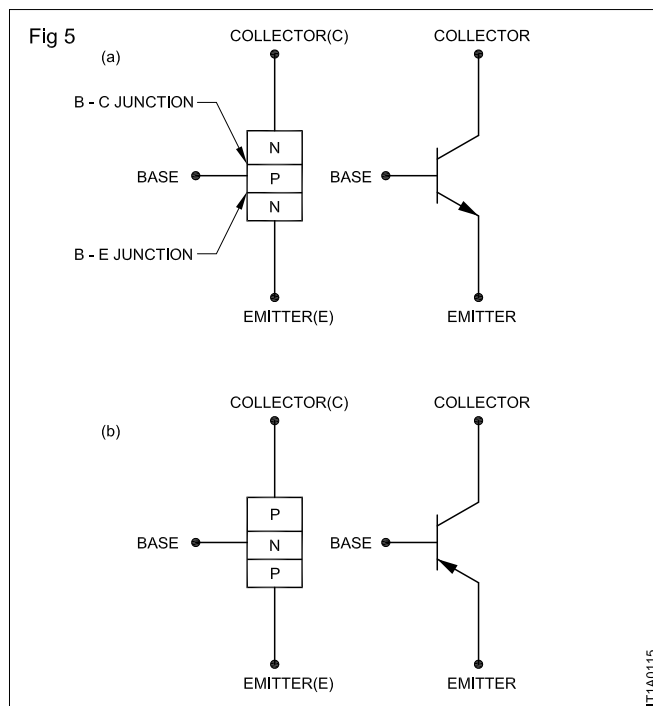
2 Based on the way the P and N junctions are organized as shown in Fig 5.

- NPN transistors
- PNP transistors

Both NPN and PNP transistors are equally useful in electronic circuits. However, NPN transistors are preferred for the reason that NPN has higher switching speed compared to PNP.

Details of switching speed is discussed in further lessons.

Whether a transistor is PNP or NPN can be found with the help of transistor data book.



3 Based on the power handling capacity of transistors as shown in Table below (Fig 6).

Low power transistors (less than 2 watts)	Medium power transistors (2 to 10 watts)	High power transistors (more than 10 watts)
Fig6 TO-92 	TO-05 	TO-03

Low power transistors, also known as small signal amplifiers, are generally used at the first stage of amplification in which the strength of the signal to be amplified is low. For example, to amplify signals from a microphone, tape head, transducers etc.,

Medium power and high power transistors, also known as *large* signal amplifiers are used for achieving medium to high power amplification. For example, signals to be given to loudspeakers etc. High power transistors are usually mounted on metal chassis or on a physically large piece of metal known as *heat sink*. The function of heat sink is to, take away the heat from the transistor and pass it to air.

Transistor data books give information about the power handling capacity of different transistor.

4 Based on the frequency of application

- Low freq. transistors (Audio frequency or A/F transistors)
- High freq. transistor (Radio frequency or R/F transistors)

Amplification required for signals of low or audio range of frequencies in Tape recorders, PA systems etc., make use of A/F transistors. Amplifications required for signals of high and very high frequencies as, in radio receivers, television receivers etc., use R/F transistors.

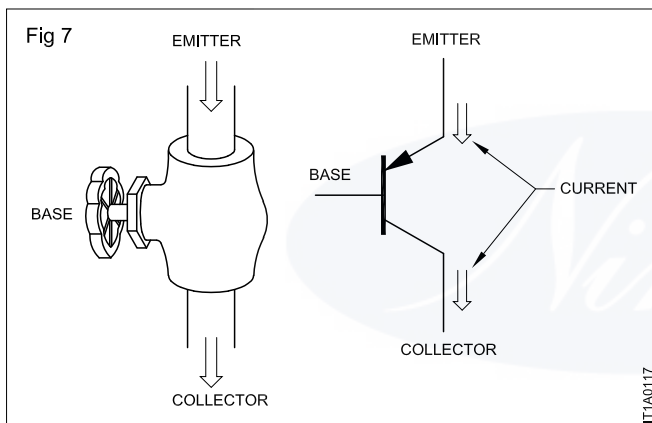
The differences between A/F and R/F transistors are discussed in detail in further lessons.

Transistor data books give information for any particular transistor as to whether it is a AF or RF transistor.

INSIDE A TRANSISTOR

Inside a transistor there are two PN junctions connected to each other as shown in Fig 3 and Fig 5. Outside a transistor, one can see only three leads. These leads are known as **base**, **emitter** and **collector** as shown in Fig 5. As shown in Fig 5, the three leads/pins/pigtails called **base**, **emitter** and **collector** are taken from each of doped semiconductor material.

In simple terms, as shown in Fig 7, the function of the **base**, **emitter** and **collector** regions of a transistor are,



Emitter - emits current carriers (electrons/holes)
 Collector - collects current carriers
 Base - controls flow of current carriers from emitter to collector.

While connecting a transistor to a circuit, it is necessary to identify the base, the emitter and the collector pins. A Transistor data book gives information on pin identification of transistors. However, it is convenient to put sleeve wires over the transistor pins for the following reasons;

- for easy identification while wiring
- sleeves act as spacers while mounting and soldering
- they ensure the required minimum lead distance from the solder joint to the transistor body.

Following colour scheme is suggested for putting sleeves to transistor pins although, any convenient colour scheme may be adopted.

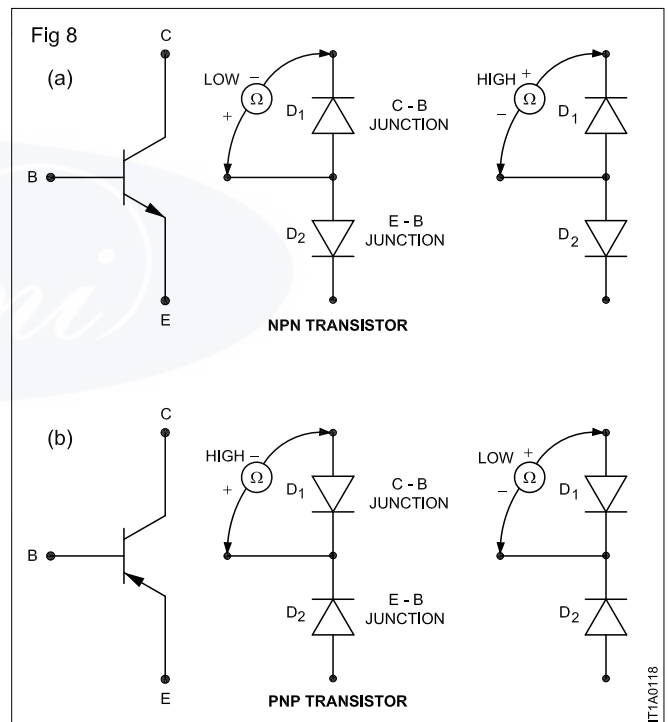
Base pin	-	Blue colour sleeve
Emitter pin	-	Red colour sleeve
Collector pin	-	Yellow colour sleeve
Shield pin	-	Black colour sleeve

TESTING TRANSISTORS USING OHMMETER

1 Junction test

Since a transistor can be regarded as two diodes connected back-to-back, a transistor's general working condition (quick-test) can be assessed by checking these two diodes as shown in Fig 8a and 8b.

Fig 8a shows a NPN transistor and Fig 8b shows a PNP transistor. The imaginary diodes 1 and 2 can be tested as testing any diode. When a diode is tested, if the ohmmeter shows high resistance in one direction and low resistance in another direction, then the diode corresponding to that diode junction can be regarded as GOOD. One important point to note in a transistor is that, both the diodes of the transistor should be GOOD to declare the transistor as GOOD.



While testing, a transistor using ohmmeter, it is suggested to use the middle ohmmeter range (Rx100) because, ohmmeters in low range can produce excessive current and ohmmeters in high range can produce excessive voltage which may be sufficient to damage small signal transistors.

2 QuickTURN-ON test

Recall that the base lead of the transistor controls the flow of current carriers from emitter to collector. So, if the base is open, then there can be no current flow through emitter-collector. This means, the resistance between emitter and collector will be high when the base is open as shown in Fig 9a. This can be checked using an ohmmeter with the base lead open.

In Fig 9, the +ve and -ve indicated across ohmmeter is the internal battery polarities of the meter and not the markings at the meter prod sockets.

When the collector and base leads of a transistor is touched with a wet finger as the base of the transistor turns ON the transistor and makes current to flow through emitter-collector. Because of the current flow, the resistance across emitter-collector will be low. From this test it is possible to make a quick test of the transistors basic operation. This test is most suitable for low power and medium power transistors.

The above two tests on a given transistor, using a simple ohmmeter reveals the condition of the transistor. These tests are essential before using a transistor in a circuit.

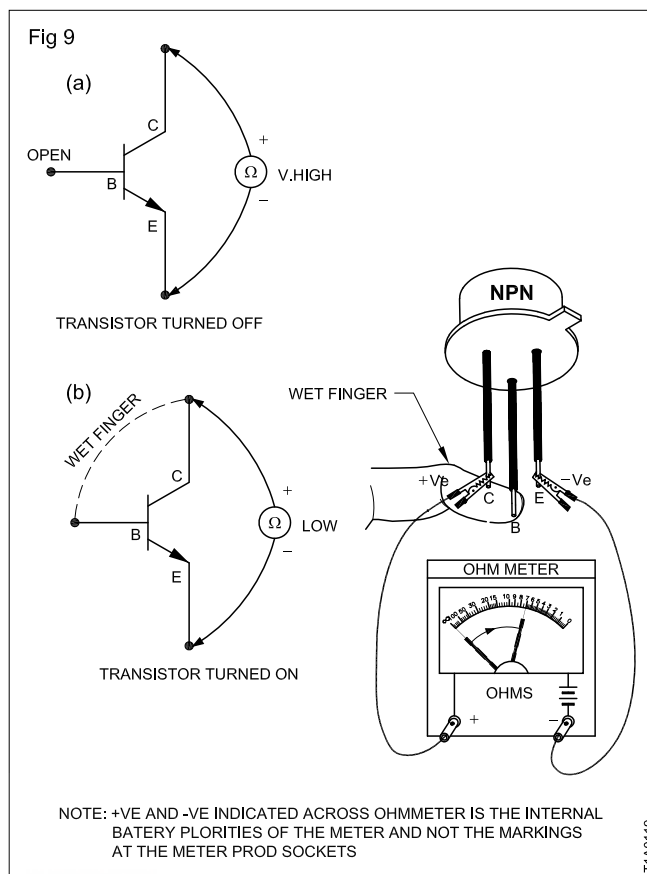
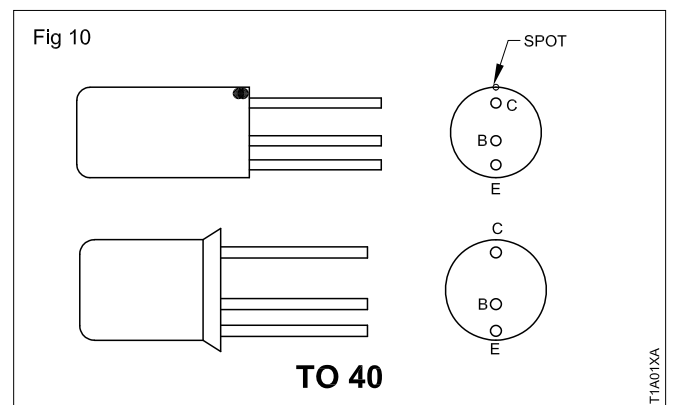
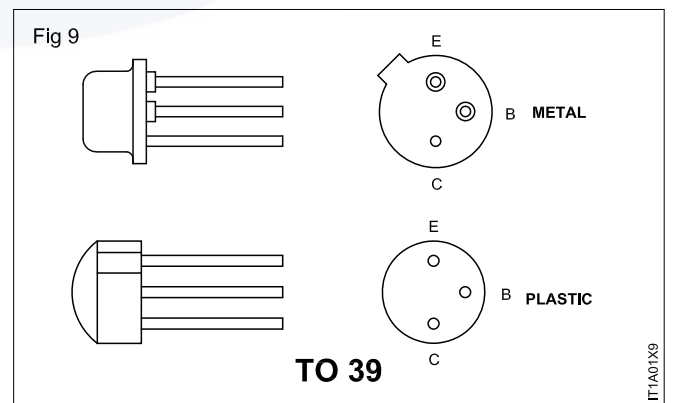
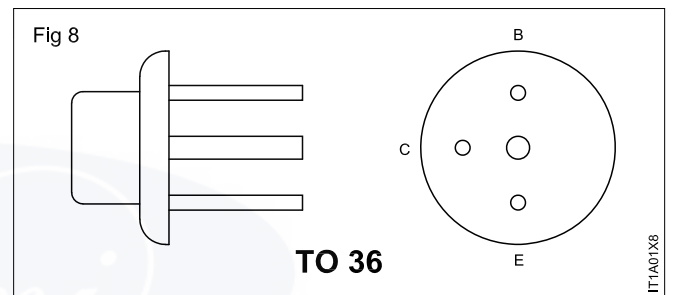
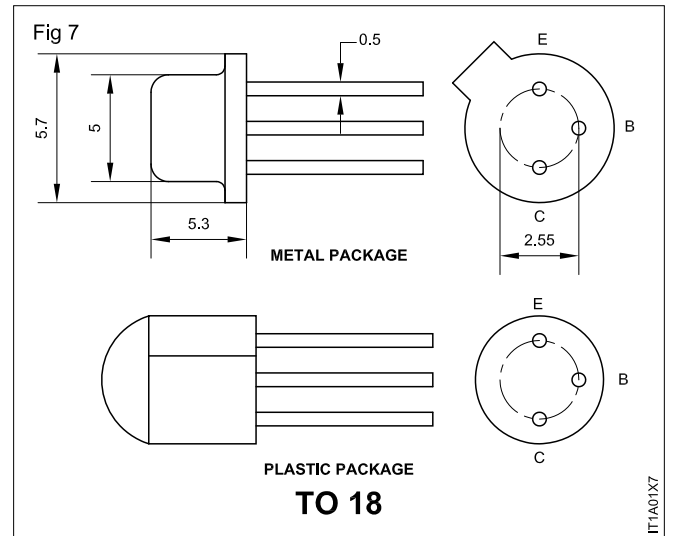
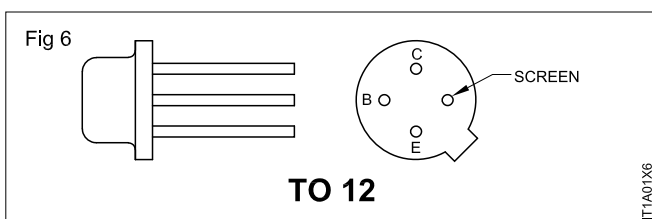
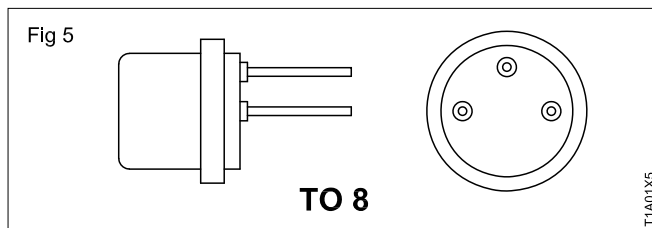
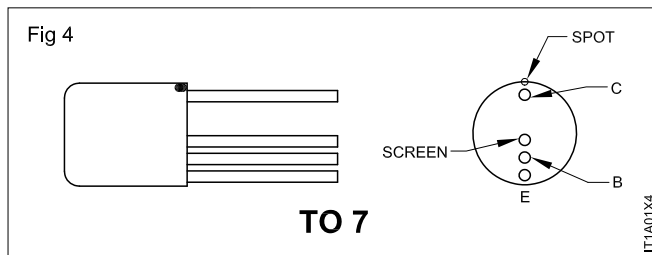
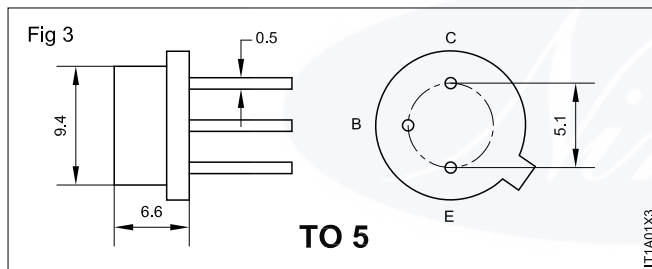
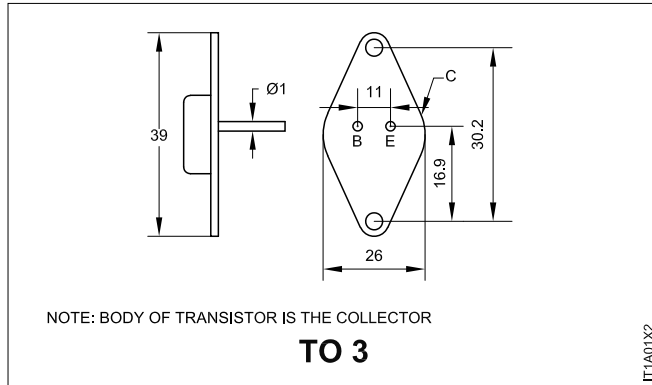
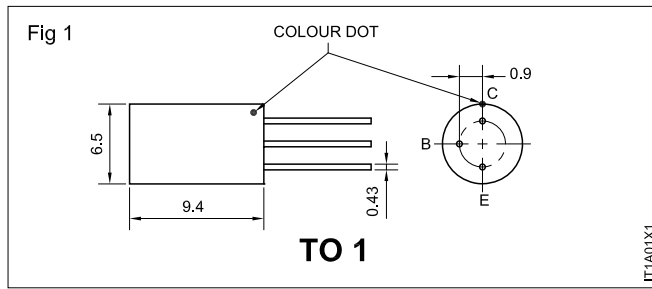
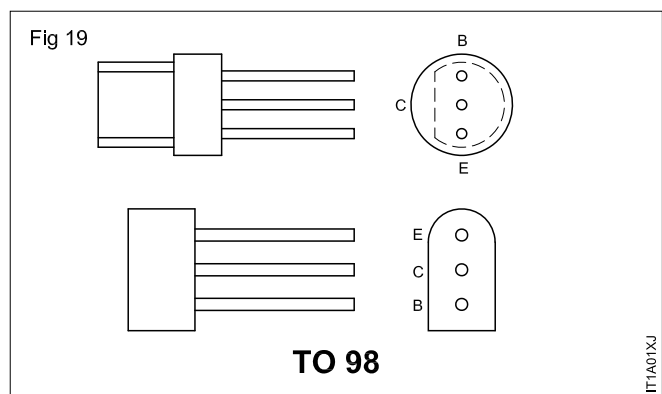
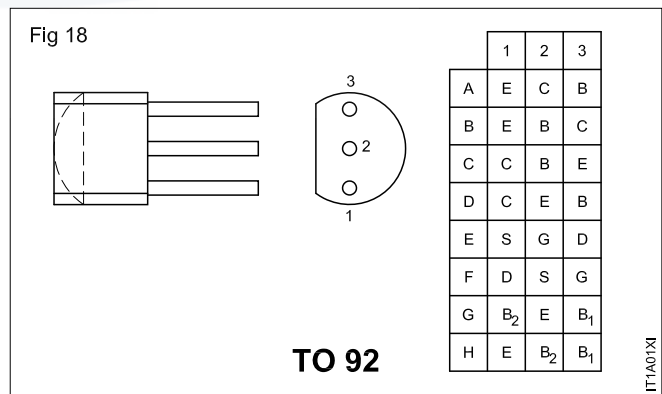
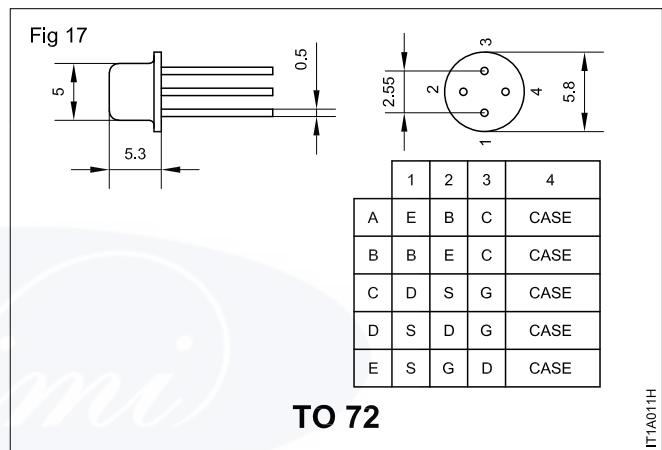
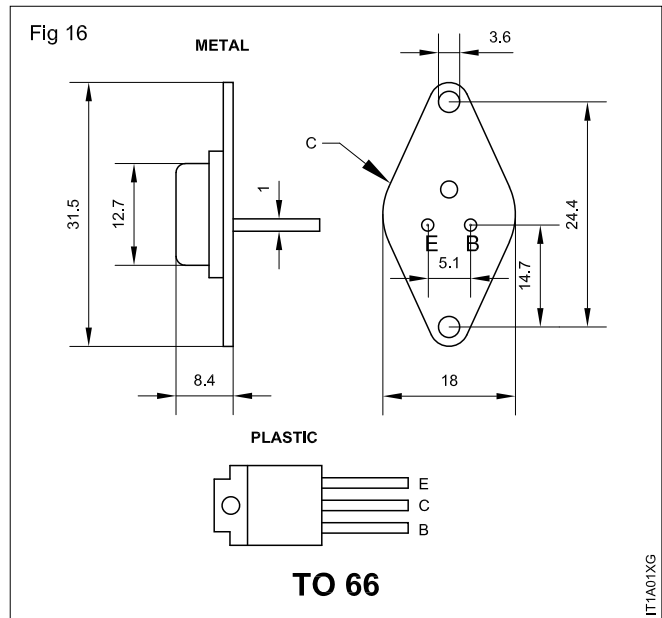
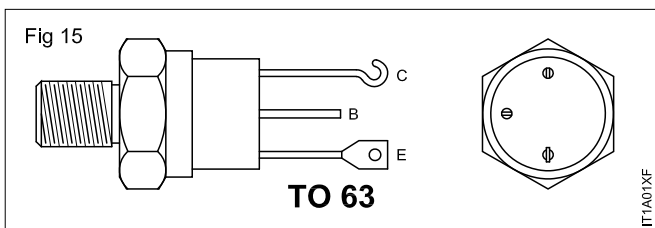
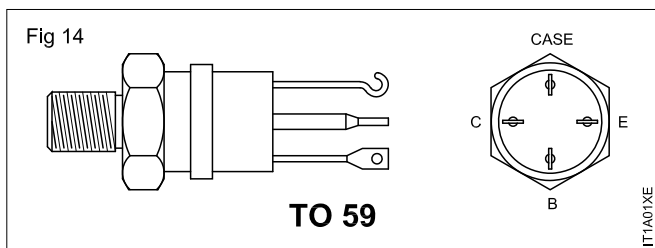
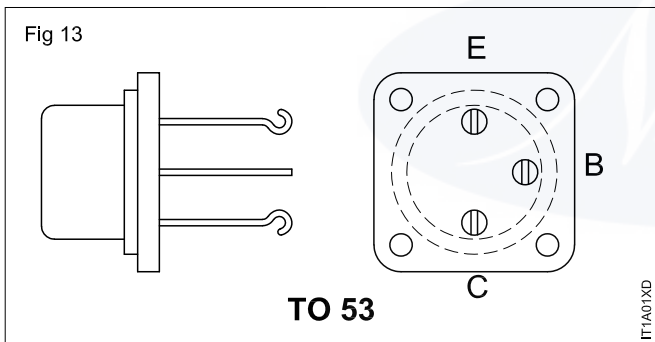
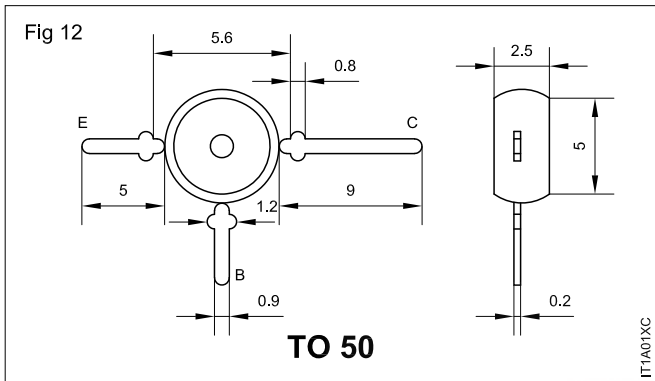
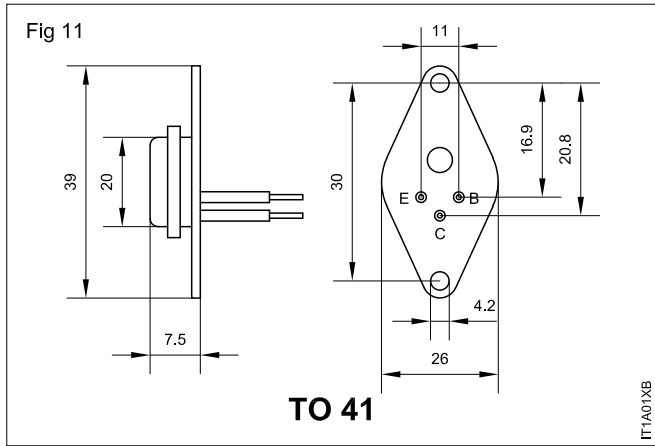
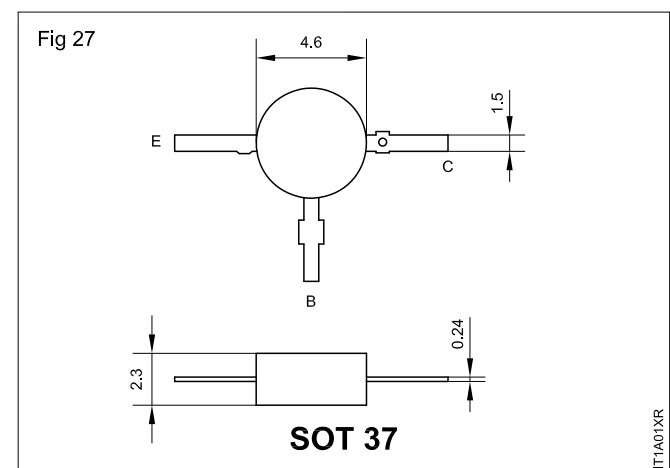
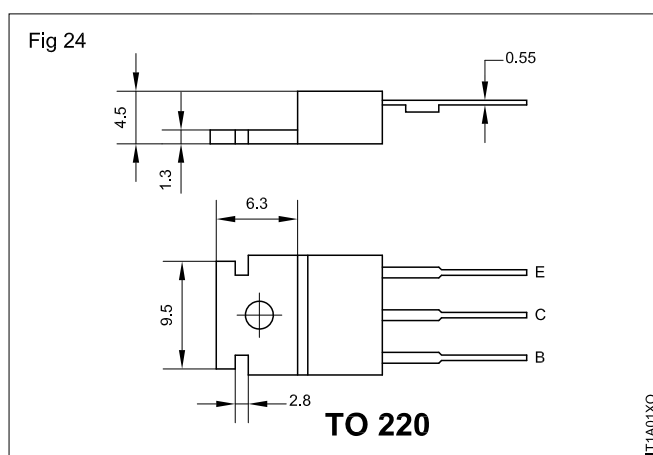
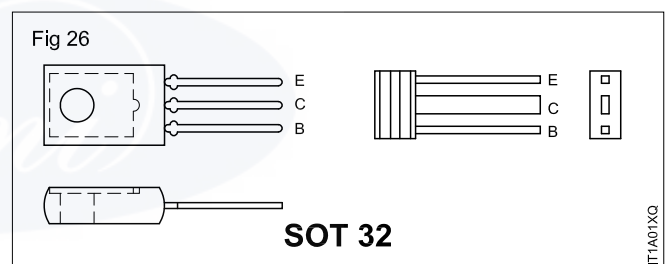
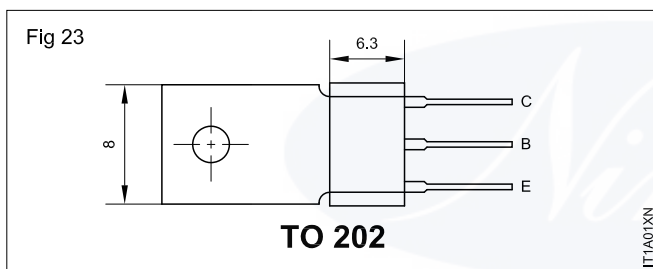
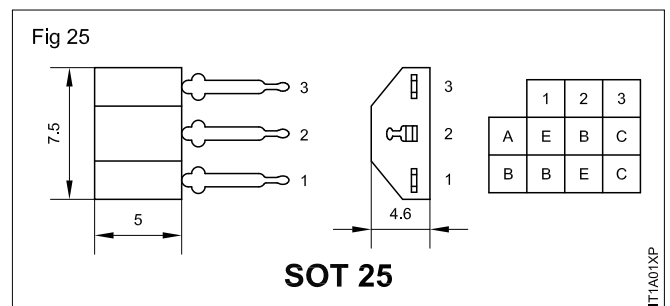
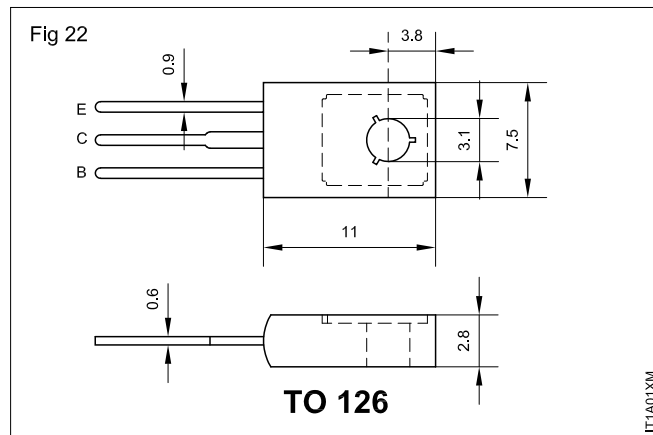
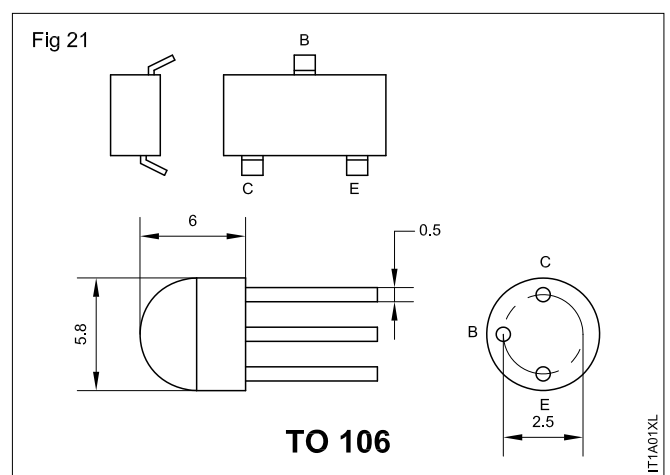
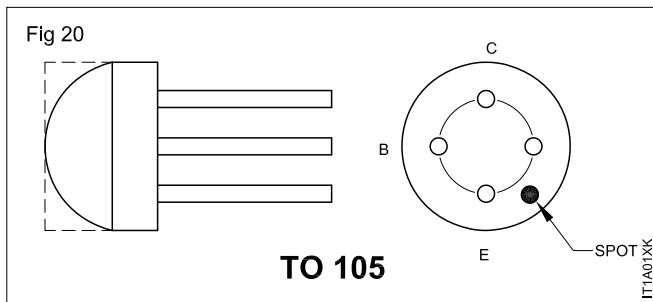


CHART - 1

Types of transistor packages and pin identification chart







Biasing of Transistors and common Emitter characteristics

Objectives: At the end of this lesson you shall be able to

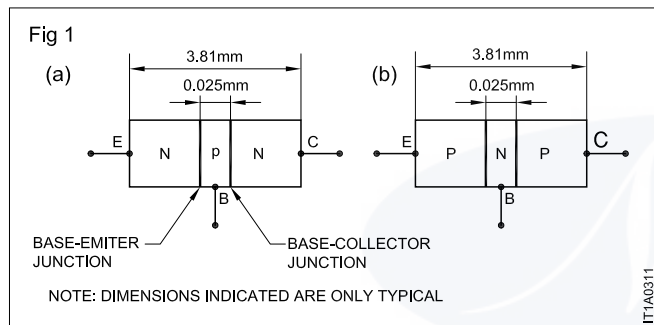
- list the important characteristic curves of transistors
- state the meaning of knee voltage and break down voltage in the output characteristic curve
- state the importance of DC load line curves
- state the meaning of a Q-point and the method of fixing Q-point for a given transistor.

Biasing of transistors

Biasing a transistor means giving correct polarity and current level of voltages at the terminals of a transistor, such that, it functions as intended. (as an amplifier or as a solid state switch etc.)

Recall, transistors are three-layer semiconductor devices consisting of either a P-type layer sandwiched between two N-type layers as shown in Fig 1a or N-type layer between two P-type layers as shown in Fig 1b.

From Fig 1, the following points are important to note;



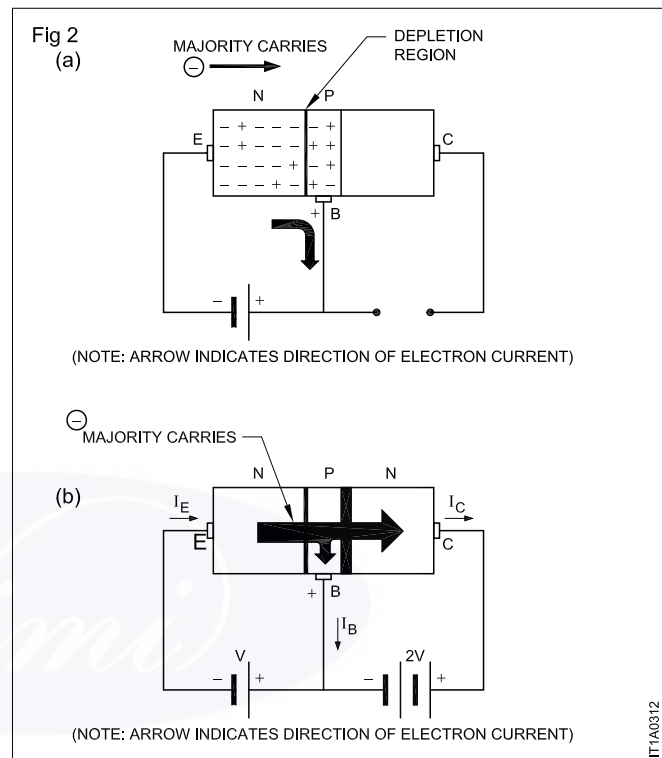
- The widths of the outer layers, i.e. emitter and collector layers are much greater than that of the base layer.
- The emitter layer is heavily doped compared to both the base and collector layers.
- The base layer is very thin, of the order of 1/10th the width of the outer layers, and is very lightly doped.

Transistor operation

As transistors have three layers, there are two junctions as shown in Fig 1. The **base-emitter** junction behaves as one diode junction. The **base-collector** junction behaves as the other diode junction.

Recall that a diode junction conducts only when +ve supply is connected to the P material and -ve supply to the N material. Fig 2a shows a NPN transistor where the base-emitter junction is forward-biased. Hence, the diode conducts resulting in large flow of majority carriers (electrons) from N-type to P-type material.

Fig 2b shows the base-emitter junction forward biased and the base-collector junction is reverse-biased. Why is the base-collector reverse biased? what effect does this connection have on the transistor operation?

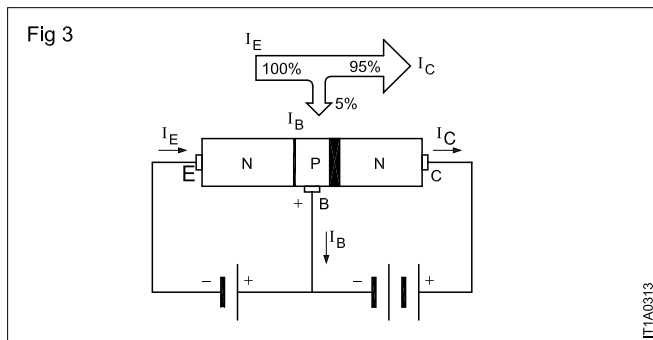


The answer is, in a NPN transistor, majority carriers are electrons, because, the emitter and collector are N-type materials. Free electrons are generated in the N-type emitter because of the forward-biased base-emitter junction. If the collector voltage is not there, then all the generated electrons flow to the base as shown in Fig 2a.

When the base-collector is reverse-biased, then, a positive voltage appears at the collector. This positive voltage at the collector completely changes the path of the electron current flow. Because of the thin base and the low base-to-emitter voltage (0.7V for silicon), about 95 percent of the electrons pass through the thin base and are attracted to the more positive potential collector as shown in Fig 2b. Only a very small percentage of the electrons from the emitter combine with holes in the base.

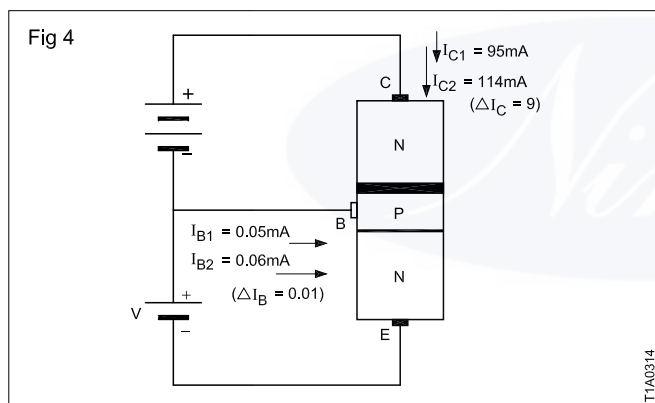
It can be seen from Fig 3, that the,

- current carriers come from the emitter
- base current is small (5% of emitter current)
- and, the collector current is high (95% of emitter current).



Under such conditions, it can be seen that, small changes in the emitter-base current will result in large change in the collector current. For example, an increase of say one electron in base current will result in an increase of 19 electrons in the collector current. This is because the collector current is 95% of the emitter current whereas the base current is only 5% of emitter current. This means that the value of the collector current can be easily controlled by changes in the bias on the emitter-base junction.

Summarizing, small changes in the base current results in large changes in the collector current as shown in Fig 4. This is nothing but amplification which is the intended function of a transistor. This behaviour of a transistor is known as Transistor action.



The ratio of the change in output current to the change in the input current is called the **amplification** or **gain**. In Fig 4, change in output current is ΔI_C due to the change in the input current ΔI_B . Therefore the current gain introduced by the transistor is,

Current gain

$$= \frac{\text{Output current change } \Delta I_C}{\text{Input current change } \Delta I_B} = \frac{9 \text{ mA}}{0.01 \text{ mA}} = 900$$

Gain is a dimension less quantity.

This condition as shown in Fig 4, in which the two junctions of the transistor are connected to such polarities of the voltage source, such that the transistor behaves as an amplifier, the transistor is then said to be properly biased or correctly biased.

Some books use the term Forward biased instead of the term properly biased.

Summarising a transistor is said to be properly biased or correctly biased or forward biased if,

- its base-emitter junction is forward biased
- and, its base-collector junction is reverse biased.

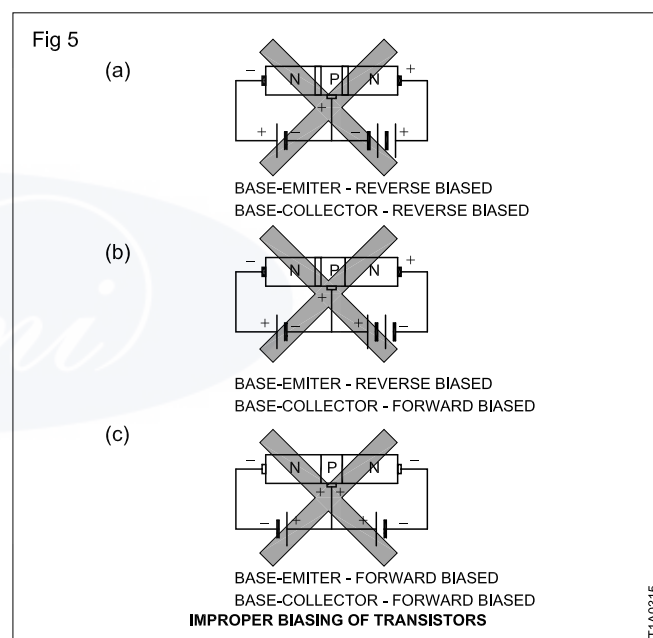
On the other hand, if the polarities of voltages connected to transistor junctions is as shown in Fig 5a and 5b, because the base-emitter junction is reverse biased, no electrons are available for conduction, and, hence, the transistor action does not exist. If the base-emitter is forward biased but the base-collector is not reverse biased as shown in Fig 5c, then, there is no amplification as both the junctions simply conduct as diodes.

In a properly biased transistor as shown in Fig 3 and Fig 4, the relationship between I_E , I_B and I_C is given by,

$$I_E = I_B + I_C \quad \dots\dots[1]$$

$$\text{or } I_C = I_E - I_B$$

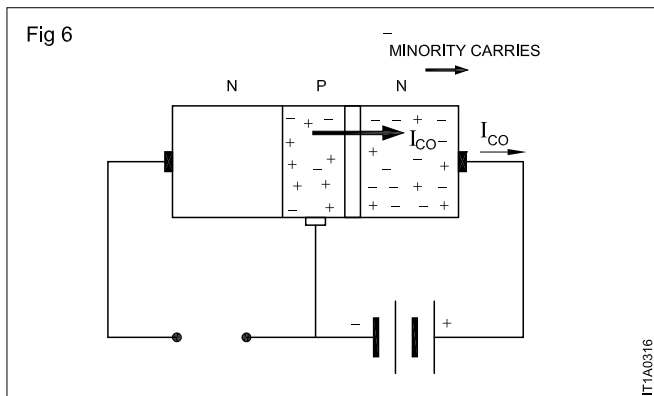
$$\text{or } I_B = I_E - I_C.$$



Minority current in transistors

In NPN transistor, as shown in Fig 6, if no voltage is applied across the base-emitter junction, but a reverse-bias is applied across the base-collector junction, the following things happen,

- There is no current in the base-emitter path as no bias voltage exists.
- The base-collector diode is reverse biased; hence, the forward current due to the majority current carriers (electrons) is zero.
- A small quantity minority current of the order of a few nanoamperes to microamperes flows in the base-collector. This small reverse current is due to minority current carriers, electrons in the **P-type** base material.

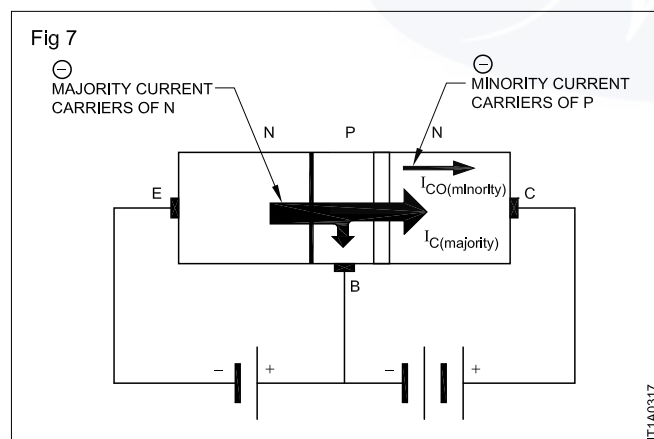


- The minority current increases if the voltage applied to the base-collector increases or the junction temperature increases. This is because current increases temperature and temperature releases current carriers from the covalent bond structure.

This minority current is called the *leakage current* and is shown by the symbol I_{CO} . I_{CO} means, collector current I_C with emitter terminal open. The value of this leakage current I_{CO} will be given in the transistor data sheets for all transistors.

As shown in Fig 7, since the forward collector current I_C due to majority carriers and the leakage current I_{CO} due to minority current carriers flow in the same direction, they are combined. Therefore, the total collector current will be equal to,

$$I_C = I_{C(\text{majority})} + I_{CO(\text{minority})} \quad \dots\dots[2]$$



Because of the minority current I_{CO} , the value of collector current I_C will be slightly more than that, given in the equation $I_C = I_E - I_B$. However, this slight increase in I_C value can be neglected because, the value of I_{CO} will be very very small compared to I_C (due to I_E) at normal working temperature.

In any typical general purpose transistor,

- value of I_C and I_E will be in milliamps
- value of I_{CO} will be in nanoamps to microamps
- value of I_B will be in microamps.

Operation of PNP transistors

Working of a PNP transistor is exactly the same as that of NPN transistors discussed earlier, if the role played by the electrons in NPN transistors is interchanged with holes as given below;

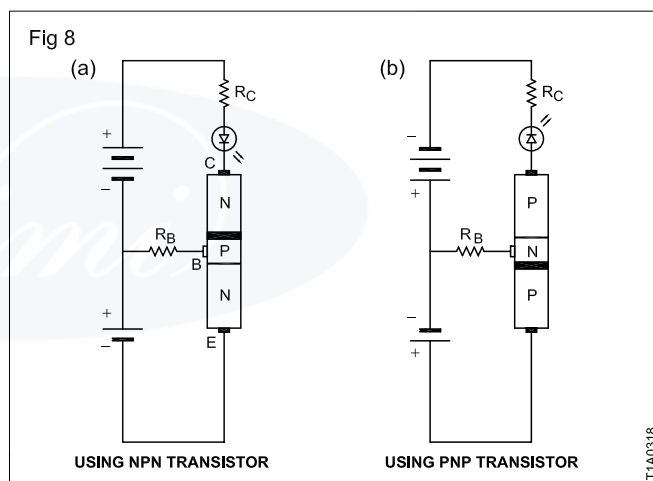
In a PNP transistor,

- the majority current carriers are holes instead of electrons
- the minority current I_{CO} , is due to electrons in the N-type base material instead of holes.

Test circuit for testing proper transistor biasing

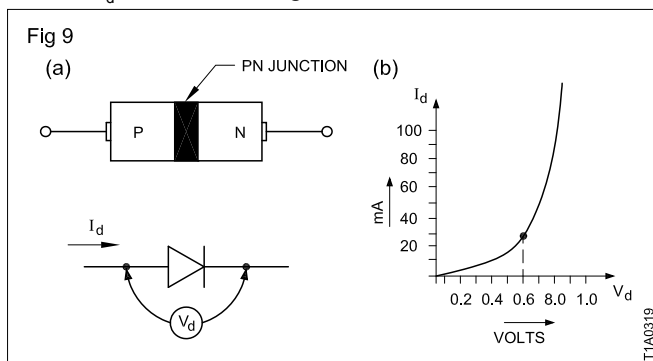
If a transistor is properly biased (i.e. B-E junction forward biased and C-B junction reverse biased), then, there will be collector current I_C of the order of milliamps. To check this an LED is connected in the collector circuit of the transistors as shown in Figs 8a and 8b. The LED in the collector glows only when the transistor is properly biased otherwise the LED remains OFF.

Resistor R_B and R_C are introduced in the circuit to limit the base and collector currents such that the transistor does not get damaged due to excessive current.

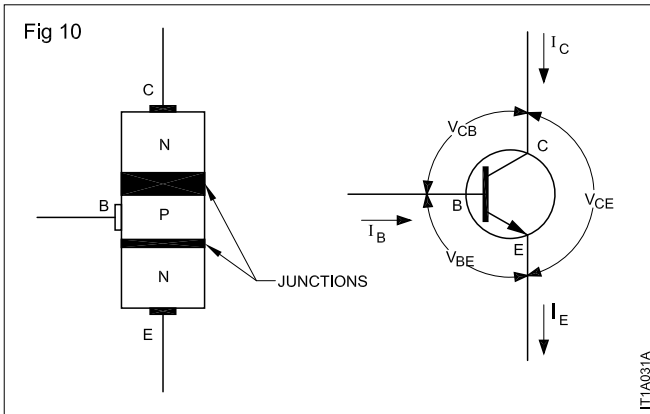


Input and output characteristics of Transistors

A semiconductor diode, as shown in Fig 9, has only one PN junction. When the voltage across the PN junction is increased or decreased, the current through the diode increases or decreases. There is only one voltage parameter (V_d) and one current parameter (I_d). Therefore, the relationship between these two parameters is easy to understand through the diode characteristic graph of V_d versus I_d as shown in Fig 9b.



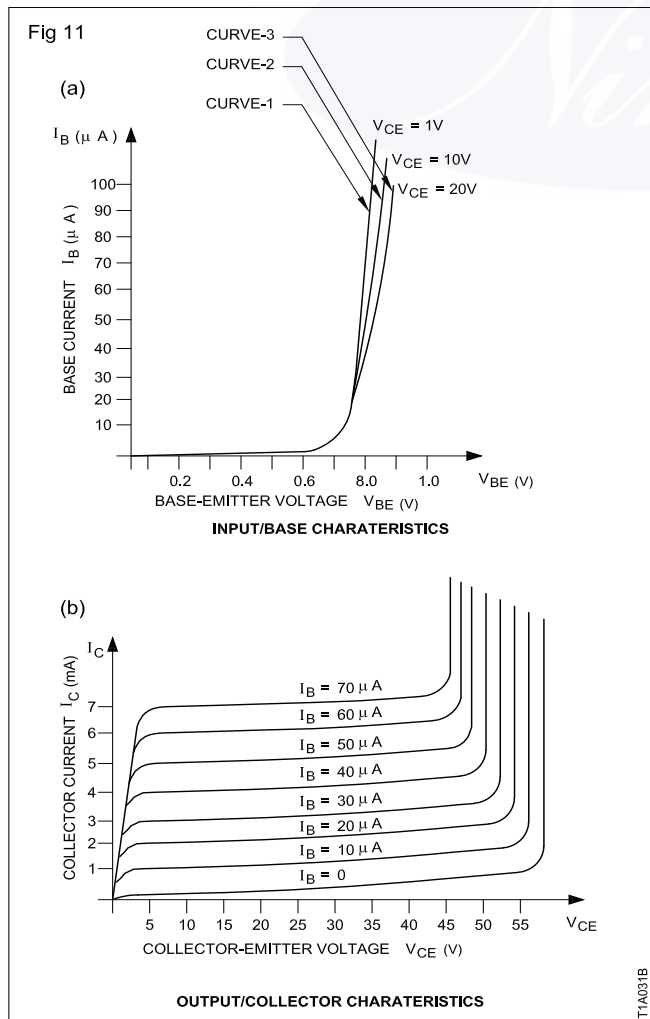
In a transistor since there are two PN junctions there are three voltage parameters V_{BE} , V_{BC} , V_{CE} and three current parameters I_B , I_C , I_E , as shown in Fig 10.



Any change in any one parameter causes changes in all the other parameters. Hence it is not very easy to correlate the effect of one parameter with the others. To have a clear understanding of their relationship a minimum of two characteristics graphs should be plotted for any transistor. They are,

- Input characteristics
- Output characteristics.

For simplicity in understanding, consider a common-emitter amplifier. The two characteristics graphs are shown in Fig 11.



Input characteristics or Base characteristics

The graph at Fig 11a shows the relationship between the input voltage V_{BE} and input current I_B for different values of V_{CE} .

Since the base-emitter section of a transistor is nothing but a diode, the graph resembles a diode curve as in Fig 9b. But, it is important to note that in Fig 11a, there is a diode curve for each value of the collector-emitter voltage V_{CE} .

While plotting the diode curve 1 of Fig 11a, the value of V_{CE} was maintained constant at 1V. In curves 2 and 3, the value of V_{CE} was increased and hence the path of the curve becomes different.

Why does this happen? The answer is, because of the higher collector voltages, the collector gathers a few more electrons flowing through the base-emitter. This reduces the base current. Hence, the curve with higher V_{CE} has slightly less base current for a given V_{BE} . This phenomenon is known as Early effect.

The gap shown between the curves in Fig 11a is very small. In practice, this gap will be so small, sometimes not even noticeable.

Plotting input/base curves of any given transistor

Necessary data for plotting the input or base characteristics can be obtained by constructing a simple test circuit as shown in Fig 1 of Exercise 13.08.

In this test circuit, V_{CE} is set to the required value by adjusting the voltage source V_{CC} . A resistor is introduced in the collector of the transistor to prevent excessive current in the collector which may damage the transistor.

The base-emitter voltage V_{BE} can be set by adjusting the potentiometer. An additional resistor is introduced in series with the dc supply V_{BB} and the POT only to limit the voltage across V_{BE} , and hence, the base current.

Output characteristics or Collector characteristics

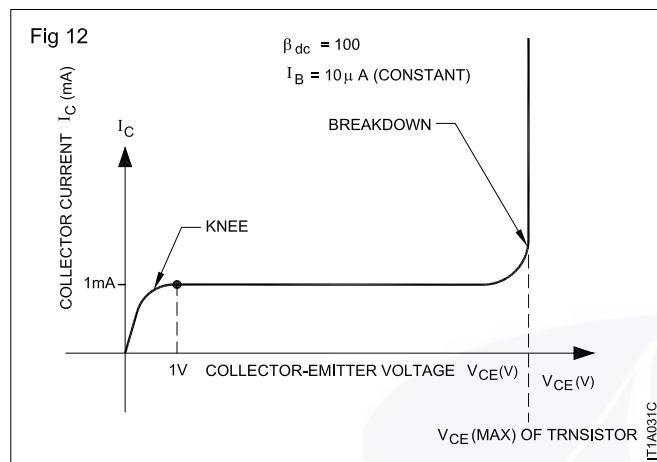
The graph at Fig 11b, shows the relationship between the output voltage V_{CE} and output current I_C for different values of I_B .

For simplicity in understanding, consider one of the curves of Fig 11b for a particular value of I_B as shown in Fig 12.

Fig 12 shows the collector characteristics for a constant I_B of $10\mu A$. Behavior of I_C for different values of V_{CE} is explained below:

- When V_{CE} is 0, the collector-base diode is not reverse-biased. Therefore, the collector current is negligibly small.
- For V_{CE} between 0.7V and 1V, the collector diode gets reverse-biased. Once reverse biased, the collector gathers all the electrons that reach its depletion layer. Hence the collector current rises sharply and then becomes almost constant.

- Above the knee voltage and below the break down voltage, the collector current does not rise steeply or the current is almost constant even if the value of V_{CE} is increased. Thus the transistor works like a *controlled constant current source* in this region.
- Assuming that the transistor has a β_{dc} of approximately 100, the collector current is approximately 100 times the base current as shown in Fig 12 (1mA is 100 times $10\mu A$).
- If V_{CE} is further increased, beyond the break down level, $V_{CE(max)}$, the collector-base diode breaks down and normal transistor action is lost. The transistor no longer acts like a current source. As the collector-base gets ruptured, the junction is shorted and hence current increases rapidly above the breakdown point as shown in Fig 12.



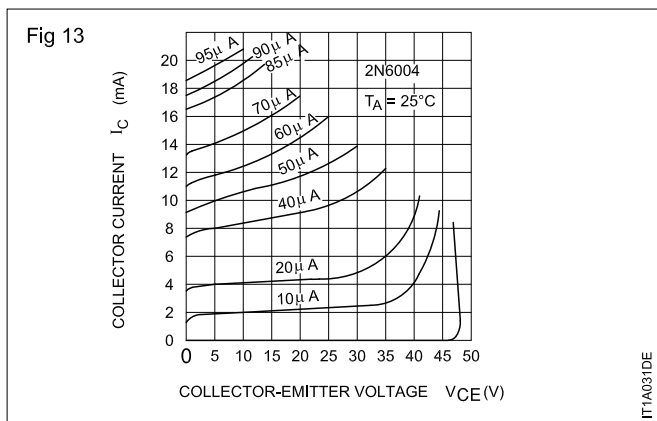
If several curves for different I_B are drawn on the same graph, the collector curves look like what is shown in Fig 11b.

Since the assumed β_{dc} of the transistor was approximately 100, the collector current is approximately 100 times greater than the base current at any point in the active region. These curves are sometimes called static collector curves because DC currents and voltage are being plotted.

In Fig 11b, notice that at the bottom most curve, even though the base current is zero, a small collector current exists. This is because of the leakage current of the collector diode. For silicon transistors this leakage current is so small that it can be almost ignored.

In Fig 11b, also note that the break down voltages become lower at higher currents. This means that the base-emitter voltage from knee point till the break down, known as the voltage compliance of a transistor, decreases for larger collector currents. Hence, it is necessary to avoid very high collector current such that the transistor operates in a wider active region.

Generally transistor data sheets/books do not show collector curves of transistors. To see the collector curves of a particular transistor, an instrument known as *curve tracer* is used. This instrument looks similar to an oscilloscope. It displays collector curves similar to those shown in Fig 13.



If different transistors are tried on the curve tracer, you will notice changes in knee voltage, β_{dc} , breakdown voltage, etc.

TIP : Two transistors of the same type number (e.g. 2N6004) may have a wide variation in the collector curves.

The collector curves are very important because, from these curves the following important information required while designing an amplifier circuit using a particular type of transistor can be obtained;

- DC current gain β of the transistor at different set DC values of I_B and V_{CE} .
- Maximum value of V_{CE} that can be applied for a set value of I_B and I_C .
- Maximum value of I_C that can be made to flow for a set value of I_B .

DC LOAD LINES of transistors

To have a further insight into how a transistor works and in what region of the collector characteristics does it work better can be found using dc load lines.

Consider a forward biased transistor as shown in Fig 14a. Fig 14b shows the collector characteristics of the transistor used.

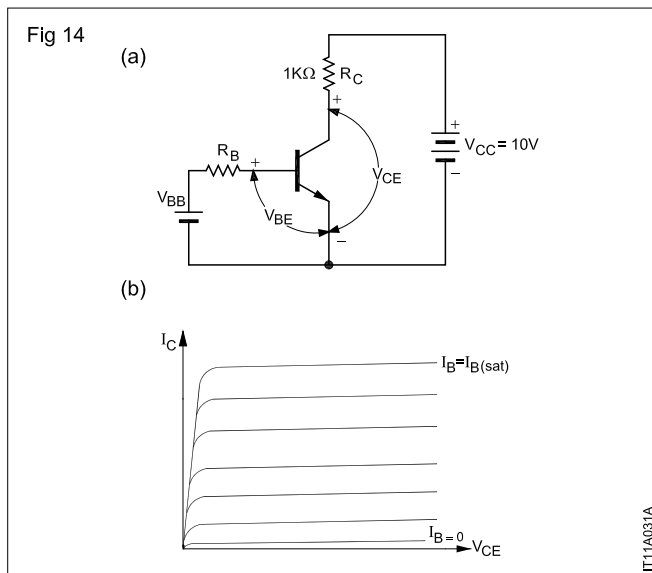
In the circuit at 14a, consider the following two situations,

- Maximum collector current, $I_{C(max)}$
- Minimum collector current, I_C

For situation (1), assume that V_{CE} is zero or collector-emitter is a short. In that case, the collector current is limited only by the collector resistor R_C .

$$\text{Therefore, } I_C = \frac{V_{CC}}{R_C} \text{ at } I_{CE} = 0$$

Under such a condition, for the circuit at Fig 14a, I_C will be equal to $10V/1KW = 10mA$.



Mark this $I_C = 10\text{mA}$ point along $V_{CE} = 0$ on the collector characteristics of the transistor as shown in Fig 15a.

For situation (2), assume that V_{CE} is maximum or collector emitter is open. In that case, the collector current is zero.

Therefore, $V_{CE} = V_{CC}$. In the circuit at 6a, $V_{CE} = V_{CC} = 10\text{V}$

Mark this point of $I_C = 0$ and $V_{CE} = 10\text{V}$ on the collector characteristics of the transistor as shown in Fig 15b.

Connect the two marked points through a straight line as shown in Fig 15c. This line is called the *load line*.

The point at which the load line intersects the $I_B = 0$ is known as the *cut off point*. At cut off, $I_B = 0$; hence, emitter diode is out of forward bias and the transistor action is lost.

The point at which the load line intersects $I_B = I_{B(sat)}$ is called the *saturation point*. At this point the base current is maximum and the collector current is also maximum. At saturation, the collector diode comes out of the reverse bias, and hence, the normal transistor action is lost.

For a transistor to work in a normal way, i.e. as a *controlled current source*, it must not be made to work either in the *cut off* or in *saturation*. Therefore the ideal point would be somewhere in the middle of these extreme points on the load line. This middle point is known as *Quiescent point* or *Q-point* as shown in Fig 15c.

Knowing the Q point, you can fix-up the value of resistors R_C and R_B of the circuit.

The dc load line shows at a glance the active V_{CE} voltage range of the transistor. In other words, it indicates that the transistor acts like a current source anywhere along the dc load line, excluding the saturation or cut off, where the current-source action of the transistor is lost.

Fixing Q point from the data available in data sheets

The Q point can be fairly accurately fixed from the data of a transistor given in transistor data books. This reduces the time consuming work of plotting the collector characteristics and the load line. To do so, the following points are very important to remember;

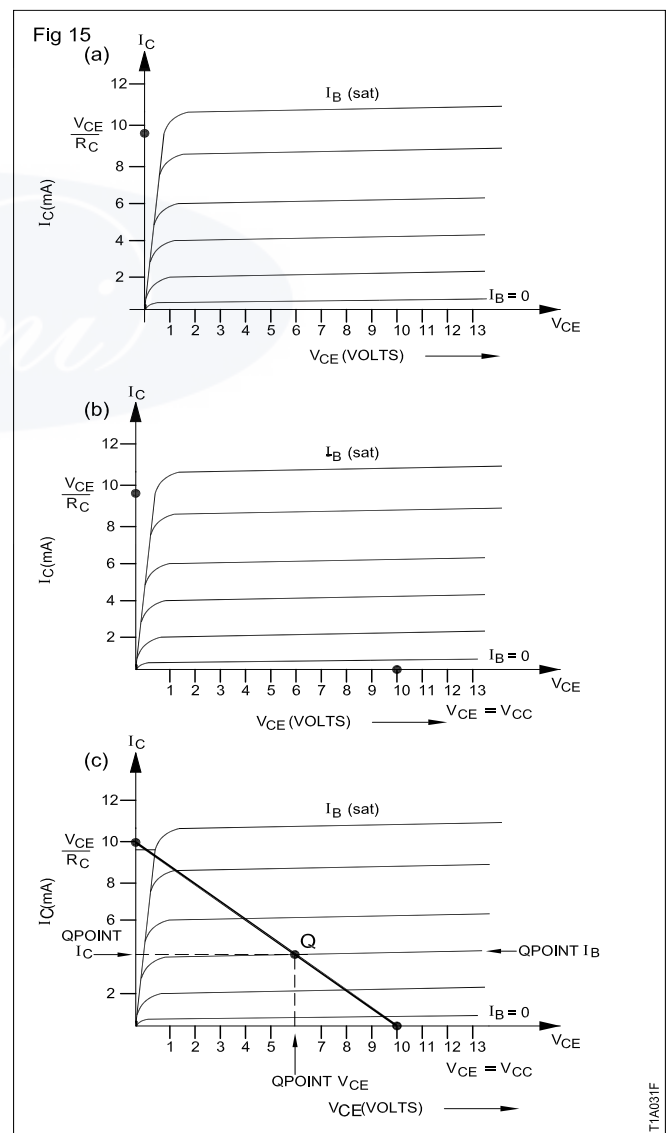
- 1 The chosen V_{CC} must be less than $V_{CE(max)}$ given in the data book
- TIP : Preferably restrict value of V_{CC} to $3/4$ of $V_{CE(max)}$
- 2 Fix the Q point I_C at $1/2$ of $I_{C(max)}$ given in the data book.
- 3 At the Q point assume $1/2$ of V_{CC} will be across V_{CE} .
- 4 From points (2)&(3) calculate the value of R_C .
- 5 From the H_{FE} value given in data book, fix the approximate value of the base current at the Q point as given below.

I_B at Q point =

Chosen value of I_C at the Q – point (tip no.2)

Typical value of H_{FE} from data book

- 6 From the value of I_B at the Q point and allowing a 0.7 volts drop across the base-emitter, calculate the value of R_B .



Gain and Impedance of common Emitter amplifier

Objectives: At the end of this lesson you shall be able to

- state the meaning and method of finding voltage gain
- state the meaning and method of finding input impedance
- state the meaning and method of finding output impedance
- state the meaning and method of finding power gain
- state the phase relationship between input and output in a CE amplifier.

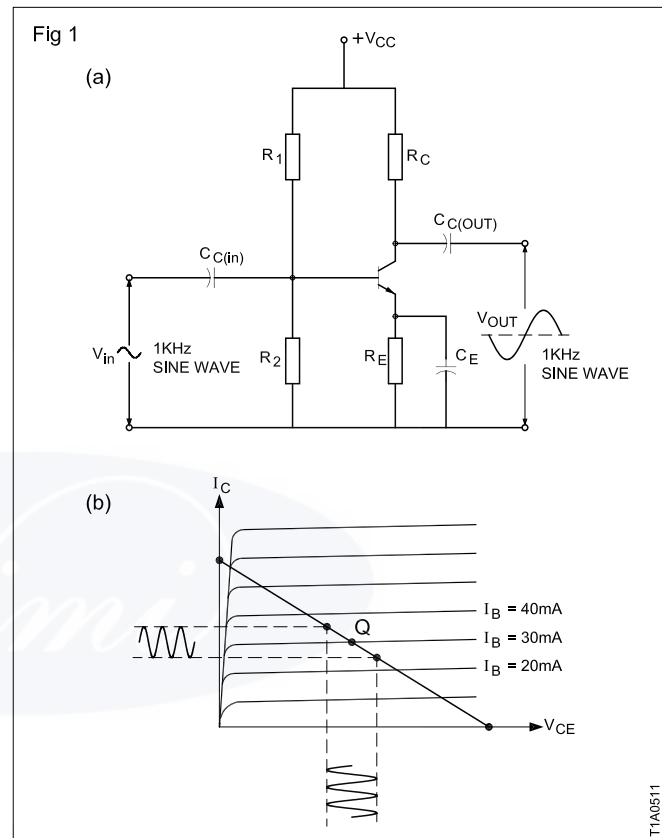
After a transistor is biased with the Q point near the middle of the DC load line, the transistor can be made to amplify ac and dc signals as shown in Fig 1a. When we use a transistor to amplify a small ac signal, the small ac signal to be amplified is coupled to the base of the transistor using a capacitor. A capacitor is used for ac coupling because as discussed in earlier lessons capacitors behave as short for ac signal and open for dc signal. The varying amplitude and frequency of the coupled ac signal produces greater value variations in the collector current of the same shape and frequency as shown in Fig 1b.

As shown in Fig 1a, if the input is a 1 KHz sine wave, the output will be an enlarged 1 KHz sine wave. The small sine wave given at the base of the transistor produces variations in the base current. Hence, the collector current is an amplified sine wave of the same frequency. The sinusoidal collector current flows through the collector resistor and produces an amplified sine wave output. Such amplifiers which retain the shape of the input signal at the output are called *linear amplifiers*.

Fig 1b, shows the dc load line, the Q point and ac input and output signals. This is generally referred to as the *ac load line*. As can be seen from Fig 1b, the ac input voltage produces variations in the base current. This results in sinusoidal variations about the Q point. Variations in Q point are nothing but the variations in the collector current resulting amplified form of the input signal.

For small input signal levels, generally referred to as small signal operation, the peak to peak swing in the collector current should be less than the $\pm 10\%$ of the collector current at Q point to keep the distortion in the amplified output within acceptable limits.

For large input signal levels, generally referred to as large-signal operation, the peak to peak swing in the collector current will be larger (more than 10%). If the swing is very large, the transistor may go into saturation and cut off. This swing into saturation and cut off will clip the positive and negative peaks of the output signal. This clipping is nothing but distortion, meaning, the output will not be an exact replicate of the input signal.



AC CURRENT GAIN A_i of a CE amplifier

The ac current gain of a CE amplifier shown in Fig 1 is the ratio of the ac component of the collector current i_c to the ac base current i_b .

$$A_i = \frac{i_c}{i_b}$$

Small letter i is used to represent ac current whose value keeps changing with time.

It is to be noted that in most linear CE amplifier circuits the current gain A_i is almost equal to β_{dc} of the transistor. Therefore the following approximation can be used for A_i .

$$A_i \approx \beta$$

In the amplifier at Fig 1, if β_{dc} of the transistor is 100, then the current gain A_i of the amplifier can be taken as 100.

VOLTAGE GAIN, A_v or A_v of CE amplifier

The voltage gain of an amplifier is the ratio of ac output voltage to the ac input voltage. This is represented as,

$$\text{Voltage gain, } A_v = \frac{V_{out}}{V_{in}}$$

Small letter v is used for voltage because it is ac voltage whose amplitude keeps changing with time.

For example, in Fig 1, if the input voltage v_{in} is 80 mV_(p-p) and the corresponding output voltage v_{out} is 7.2 V_(p-p), then the voltage gain A_v is given by,

$$\text{Voltage gain, } A_v = \frac{7.2 \text{ V}_{(p-p)}}{80 \text{ mV}_{(p-p)}} = 90$$

A voltage gain of 90 means that, in this amplifier, a base voltage of 1 mV produces an output voltage of 90 mV.

The input and output voltage may be rms, peak, peak-to-peak, as long as the input and output are measured the same way consistently.

INPUT IMPEDANCE, Z_{in} of CE amplifier

Recall that the maximum transfer of power takes place when the impedances of the supplying and receiving circuits are matched.

If impedances are to be matched for best circuit operation, both impedances must be known. If a single device such as a microphone, speaker, relay, etc. is to be used, its impedance will be given by the manufacturer. The amplifier to be designed for such a circuit must have an input or output impedance to match the input-output devices.

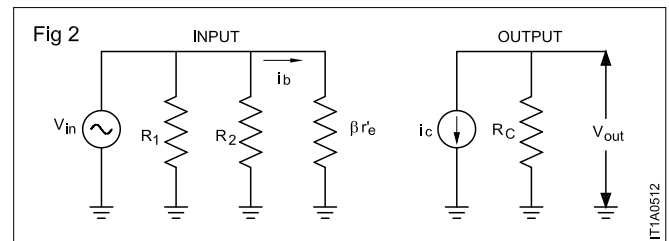
The ac source driving the amplifier has to supply ac current to the amplifier. The less the current the amplifier draws from the source, the better because the supplying source does not get loaded. The input impedance of the amplifier determines how much of current the amplifier takes from the ac source or the preceding stage of the amplifier.

In the normal frequency range of an amplifier, the coupling and by pass capacitors behave as a short for ac. The ac *input impedance* Z_{in} sometimes referred to as input resistance R_{in} is defined as the ratio of input signal voltage to input signal current.

$$Z_{in} = \frac{V_{in}}{i_{in}}$$

where, V_{in} and i_{in} are rms or peak or peak-to-peak values.

Fig 2 shows the ac equivalent circuit of the CE amplifier shown in Fig 1.



From the ac equivalent circuit the input impedance Z_{in} is given by,

$$Z_{in} \approx R_1 \parallel R_2 \parallel \beta r'_e \quad \dots [1]$$

where,

R_1 and R_2 are the voltage divider resistors,

β is the dc current gain and r'_e is the ac emitter resistance (V_{BE}/I_E). r'_e is approximately equal to 25V when the Q point is chosen at the mid of the load line.

In the CE amplifier at Fig 1, if $R_1 = 18 \text{ K}\Omega$, $R_2 = 8.2 \text{ K}\Omega$ and the transistor β is 100, the input impedance Z_{in} will be,

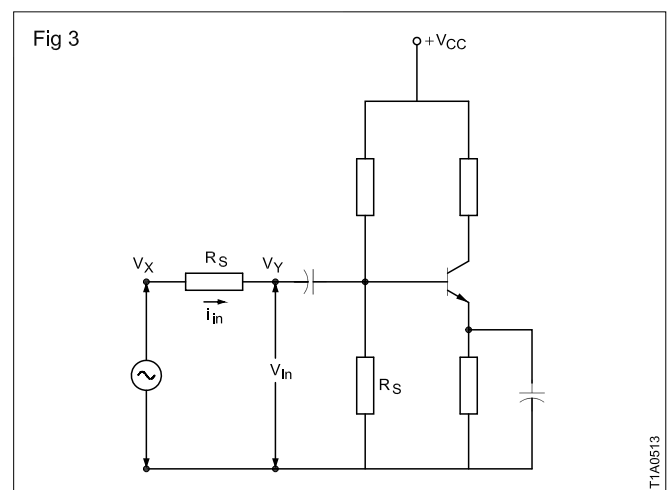
$$\beta r'_e = 100(25 \Omega) = 2.5 \text{ K}\Omega$$

$$\begin{aligned} Z_{in} &= R_1 \parallel R_2 \parallel \beta r'_e \\ &= 18 \text{ K}\Omega \parallel 8.2 \text{ K}\Omega \parallel 2.5 \text{ K}\Omega \\ &= 1.73 \text{ K}\Omega. \end{aligned}$$

Practical way of finding Z_{in}

To find Z_{in} of a given CE amplifier circuit, it is merely necessary to measure the ac signal input voltage and current. Then, use these values in the formula, and calculate Z_{in} .

A simpler method to measure i_{in} is to connect a series input resistance of known value in series with the input signal, as in Fig 3.



The voltage drop across the resistor R_s is measured, and Ohm's law is used to determine i_{in} .

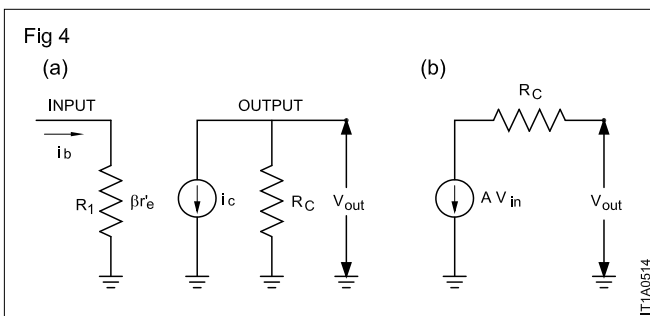
$$i_{in} = \frac{V_x - V_y}{R_s}$$

The value of V_{in} can be measured directly, as shown in Fig 3.

OUTPUT IMPEDANCE, Z_{out}

The output impedance of a CE amplifier is naturally the impedance at the output terminals.

To find the Z_{out} of the CE amplifier shown in Fig 1, consider the ac equivalent of the output as shown in Fig 4a.



Recall that a transistor operating in the linear portion of its characteristics curve is like a current source. Therefore, we can represent it as a current source i_c .

As can be seen from Fig 4a, this collector current source is in parallel with the collector resistor R_C . Assuming that the collector current source is ideal, it has infinite internal impedance. Then, the only impedance in the output is the collector resistor R_C .

The Thevenin's voltage appearing at the output is the voltage gain(A) times the input v_{in} .

Therefore, $V_{out} = A \cdot v_{in}$

Hence, the output ac equivalent circuit of the amplifier can be simplified as shown in Fig 4b. In Fig 4b, an ideal output voltage source $A \cdot v_{in}$ with zero internal impedance is in series with the collector resistor R_C . Therefore, the output impedance of the CE amplifier is approximately equal to the collector resistor R_C ,

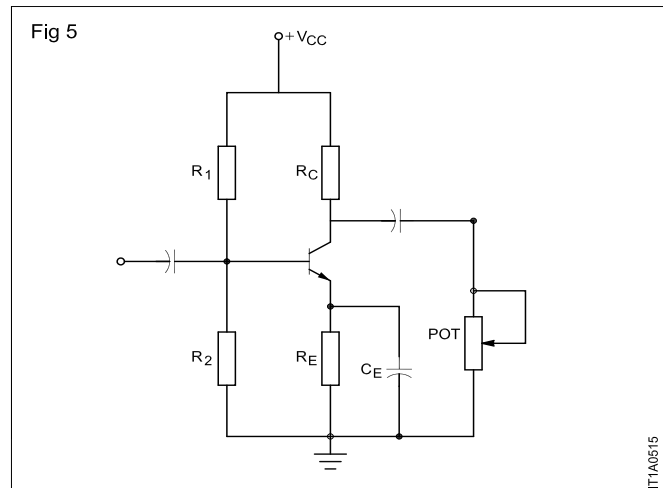
$$Z_{out} \approx R_C$$

In the CE amplifier circuit at Fig 1, if $R_C = 1000\Omega$, the output impedance of the amplifier is equal to the value of R_C , that is 1000Ω .

Practical way of finding Z_{out}

The easiest way of measuring the output impedance of a CE amplifier circuit is given below;

- (1) Measure the unloaded output voltage V_{out} of the CE amplifier.
- (2) Place a variable resistor across the load terminals, as shown in Fig 5.
- (3) Adjust the variable resistor until the voltage drop across it is one-half of the unloaded output voltage V_{out} .
- (4) Remove the variable resistor and measure its value. This value is equal to Z_{out} .



Note that Z_{out} is not a fixed value; it varies with transistor voltages and the load resistance. Care must always be taken to maintain an undistorted signal when input or output impedances are measured.

POWER GAIN, A_p of a CE amplifier

In the CE amplifier shown in Fig 1, the input power is given by,

$$P_{in} = v_{in} \cdot i_{in}$$

and the output power is given by,

$$P_{out} = -v_{out} \cdot i_{out} \cdot A_p$$

The negative sign associated with output power. This is because, in a CE amplifier, the output is 180° out of phase with the input signal. Details are discussed in the forthcoming paragraphs.

In the CE amplifier at Fig 1, power gain A_p is the ratio of output signal power to input signal power. The formula is,

$$\text{Power gain} = \frac{P_{out}}{P_{in}}$$

Power gain is also given by,

$$A_p = -A_v \cdot A_i$$

where,

A_v is the voltage gain (v_{out}/v_{in})

A_i is the current gain (i_c/i_b)

For the amplifier at Fig 1, if $A_v = 90$ and the β of the transistor is 100, then the power gain A_p of the amplifier is given by,

$$A_p = -A_v \cdot A_i = 90 \times 100 = 9000.$$

This means that if an ac input power of $1 \mu\text{W}$ is given to the amplifier, the output power will be 9mwatts.

Practical way of finding A_p

Since the formula for power is, $P = I^2 \times R = |I| \times |R|$

Since,

$$I = \frac{V}{R} \text{ (substituting this in above equation, we get)}$$

$$P = \frac{V^2}{R}$$

Therefore, by Ohm's law, power gain is easy to calculate when signal voltages and impedances are known as given below;

$$P_{out} = \frac{V_{out}^2}{Z_{out}} \text{ and } P_{in} = \frac{V_{in}^2}{Z_{in}}.$$

Knowing the values of P_{out} and P_{in} power gain of the circuit can be calculated.

Power gain, A_p in decibels, dB

The power gain of amplifiers is often expressed in decibels (dB). To calculate the power gain of an amplifier in decibels, use the following formula.

$$\text{Power gain in dB} = 10 \log \frac{P_{out}}{P_{in}}.$$

INPUT - OUTPUT phase relationship

Recall, that while calculating the power gain it was mentioned that the output signal of a CE amplifier is 180° out of phase with its input signal. To find out why this happens in a CE amplifier, assume that the dc base bias current I_b at the set Q point is $30 \mu\text{A}$. The corresponding collector current is 1 mA. When the ac signal is applied to the input, the base bias varies from 20 to $40 \mu\text{A}$, as shown in Fig 1b. Since the type of transistor used is NPN, as base bias is increased to $40 \mu\text{A}$, collector current i_c increases. The resultant effects are,

- the increased transistor conduction causes less voltage drop across the transistor (V_{CE})
- increased i_c causes a larger voltage drop across R_C . Hence, the voltage across the collector to ground gets reduced.

In Fig 1a, as the output signal is taken across the transistor collector and ground, an increasing signal voltage causes a decreasing output signal.

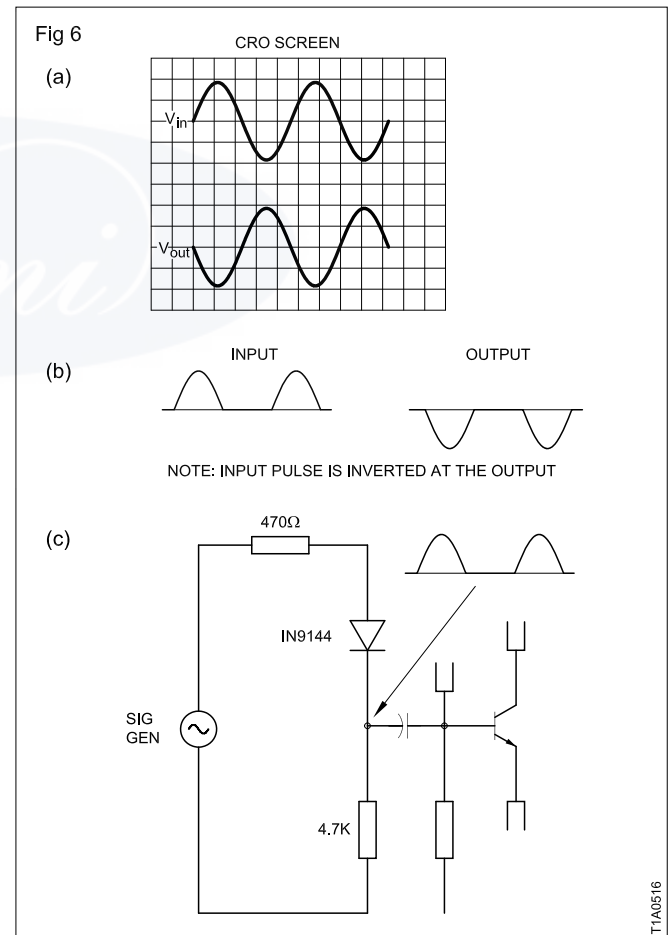
As the input signal level decreases, say to $20 \mu\text{A}$, the forward bias is less and transistor conduction decreases. When transistor conduction decreases, its resistance is higher and so the voltage drop across it increases. With increased voltage drop across the transistor, the output voltage V_{out} increases. This increase in V_{out} reduces the voltage drop across the collector load resistance R_C .

From this, it can be concluded that in a CE amplifier, a negative-going input signal causes a higher, or, more positive-going output signal. Therefore, in a CE amplifier the output is 180° out of phase with the input.

Practical way of finding input-output phase relationships

The phase relationship between input-output can be found in two ways as given below.

Method 1: Using a dual trace CRO, connect one of the CRO inputs to the input of the amplifier and the other CRO input to the output of the amplifier. Make the oscilloscope to trigger on the input signal. The waveforms shown on the CRO show the phase relationship between input and output as shown in Fig 6a.



Method 2: If a single trace CRO is used, then instead of feeding complete sinewave to the input, feed only the positive signal pulses as shown in Fig 6b. These positive pulse can be generated using a simple half-wave rectifier as shown in Fig 6c.

With the positive pulse fed at the input, the output of the CE amplifier will be negative pulse as shown in Fig 6b. Use the signal as an external trigger source for the CRO to view the waveform.

The Appendix-D gives a summary of the important parameters of a common-emitter amplifier with voltage divider bias.

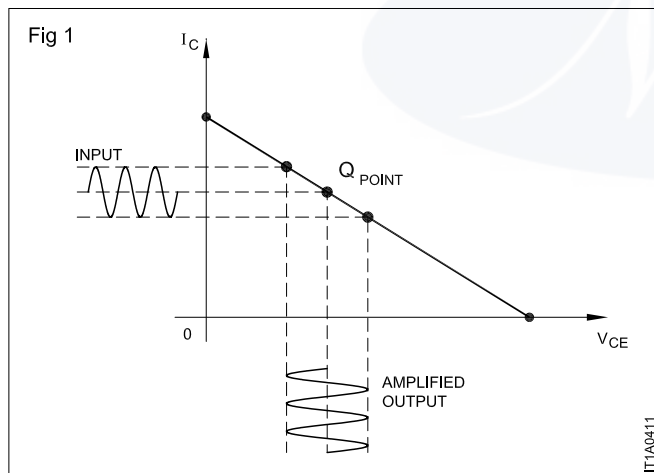
Types of Transistor biasing

Objectives: At the end of this lesson you shall be able to

- state the need for biasing of transistors
- state the reason for shifting of Q due to temperature and b_{dc} changes
- name the three main types of transistor biasing
- state the reason why base-bias is not preferred
- state the meaning of feedback
- state the reason why emitter-bias is also known as emitter feedback bias
- state why voltage divider bias is the most preferred type of biasing
- state the design guidelines for voltage divider type biasing.

Before any one rides a motor cycle or drives a car, he has to start the engine and keep the engine running. In simple terms biasing transistors is similar to keeping the transistor started before making the real use of it. Once the transistor is started, like the engine of a car, it can be made to amplify, like covering the distance by riding the car.

Before an ac signal is fed to a transistor, it is necessary to set up an operating point or the quiescent(Q) point of operation. Generally this Q point is set at the middle of the DC load line. Once the Q point is set, then the incoming ac signals can produce fluctuations above and below this Q point as shown in Fig 1.



The need for biasing of a transistor can also be explained as follows;

For a transistor to remain operating in the linear region, the emitter diode must remain forward biased and the collector diode must remain reverse biased as long as the amplifier is amplifying. In other words, the amplitude variations in current and voltage of the input signal must not drive the transistor either into saturation or cut off.

Stable Q point

A set Q point of a transistor amplifier may shift due to increased temperature and transistor b value changes. Therefore, the objective of good biasing is to limit this shifting of the Q point or to achieve a stable Q point.

Recall that, the Q point is nothing but a point in the output characteristic of the transistor. This point corresponds to a particular value of I_B , I_C and V_{CE} . Also recall that, the collector current I_C depends both on I_B and β of the transistor. If I_B changes, I_C changes, and hence, the Q point changes. If β changes, again I_C changes, and hence, the Q point gets shifted.

Shifting of Q due to temperature

Remember that a transistor is a temperature sensitive device. Any increase in the junction temperature results in leakage current. This increased leakage current in turn increases the temperature and the effect is cumulative. This chain reaction is called *thermal run away*. If this thermal run away is not stopped, it may result in the complete destruction of the transistor due to excessive heat. In transistors, due to this increased leakage current, the base current increases, and hence, the Q point gets shifted. This change in the set Q point affects the performance of the amplifier resulting in distortion.

Shifting of Q due to b_{dc} changes

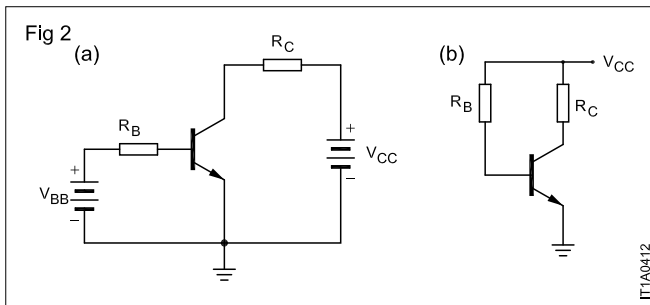
Recall that two transistors of the same type number may have different values of β . This is due to the manufacturing process of transistors. Hence, when a transistor is replaced or changed, due to different β of the replaced transistor, the Q point again gets shifted.

Therefore, a stable biasing is one which does not shift the Q-point even if temperature varies and/or the β of the transistor changes.

Types of transistor biasing

There are several ways to bias a transistor for linear operation. This means, there are several ways of setting up a Q point near the middle of the dc load line. Important biasing arrangements used with transistors are explained below:

- 1 **BASE BIAS:** Fig 2 shows one type of biasing of transistor known as base-bias. As shown in Fig 2b, usually, the collector voltage supply itself is used for the base voltage instead of a separate supply.



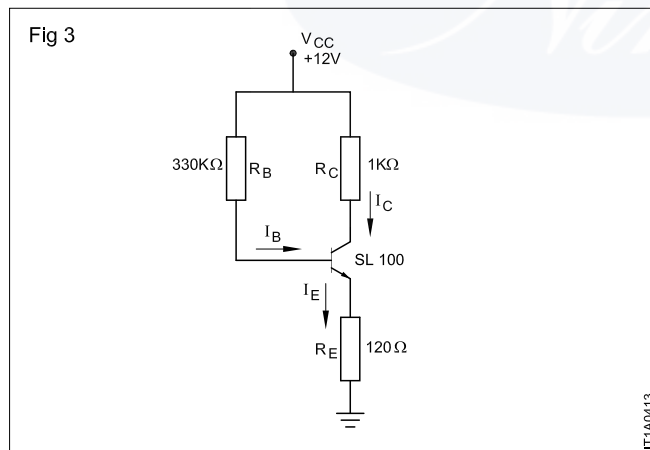
The value of the base resistor R_B is fixed such that it allows the necessary Q point base current I_B . The value of R_B ensures that the base-emitter diode is always forward biased by allowing 0.7V (for silicon) across V_{EB} .

This type of biasing is the simplest of all. However, this is the worst possible way to bias a transistor because the dc Q point changes when,

- temperature increases and
- β of the transistor is changed.

Hence, in a base-biased transistor, it is impossible to set up a stable Q point. Therefore, base biasing of transistors is not generally done in linear amplifier circuits. However, base biasing is commonly used in digital circuits (discussed in further lessons) where transistors are used as a switch and not as a linear amplifier.

2 EMITTER BIAS or emitter feedback bias: Fig 3 shows a emitter-biased transistor. This type of biasing compensates for the variations in β_{dc} and keeps the Q point fairly stable.



In Fig 3, if β_{dc} increases, the collector current increases. This in turn increases the voltage at the emitter. This increased emitter voltage decreases the voltage across the base-emitter junction and therefore, the base current reduces. This reduced base current results in less collector current, which partially offsets the increase in I_C due to increased β_{dc} .

Emitter bias is also referred to as *emitter feedback bias*. This is because an output quantity, i.e., the collector current, produces a change in an input quantity i.e., the base current. The term feedback means a portion of the output is given back to the input. In emitter bias, the emitter resistor is the feedback element because it is common to both the output and input circuits.

In Fig 3, if we add the voltages around the collector loop, we get,

$$I_C R_C + V_{CE} + I_E R_E - V_{CC} = 0 \quad \dots\dots\{1\}$$

Since I_E approximately equals I_C , (as I_B is comparatively very small), equation..(1) can be arranged as,

$$I_C = \frac{V_{CC} - V_{CE}}{R_C + R_E} \quad \dots\dots\{2\}$$

If we add voltages around the base loop, we get,

$$I_B R_B + V_{BE} + I_E R_E - V_{CC} = 0. \quad \dots\dots\{3\}$$

Since $I_E \gg I_C$ and $I_B = I_C / \beta_{dc}$, we can rewrite the equation as,

$$I_C = \frac{V_{CC} - V_{BE}}{R_E + R_B / \beta_{dc}} \quad \dots\dots\{4\}$$

From equation ... (4), the presence of term β indicates that I_C is dependent on β . The intention of emitter-feedback bias to swamp out the effect β_{dc} . This is possible when R_E is made much larger than R_B / β_{dc} . However, in practical circuits R_E cannot be made very large because, large value of R_E takes the transistor out of the linear operating region. Due to this problem, the emitter-feedback bias is almost as sensitive to changes in β_{dc} as is the base-bias. Therefore, emitter-feedback bias is also not a preferred form of transistor bias and should be avoided.

In emitter-bias, the saturation current will be,

$$I_{C(sat)} = \frac{V_{CC} - V_{CE(sat)}}{R_E + R_C} \quad \dots\dots\{1\}$$

When the transistor is saturated, the value of V_{CE} will be between 0.2 to 0.3V. Hence can be neglected for all practical purposes.

In Fig 3, the saturation current is,

$$I_{C(sat)} = \frac{12 \text{ V}}{1000 \Omega + 120 \Omega} = 10.71 \text{ mA.}$$

NOTE: $V_{CE(sat)}$ of 0.2 volts is neglected.

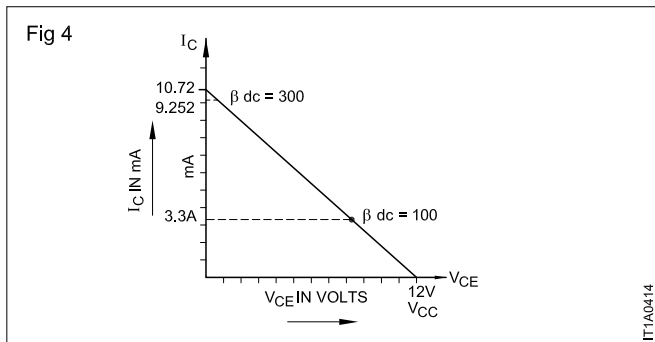
When $\beta_{dc} = 100$, equation ... (4) gives,

$$I_C = \frac{12 \text{ V} - 0.7 \text{ V}}{120 \Omega + 330 \text{ K}\Omega / 100} = 3.3 \text{ mA.}$$

When $\beta_{dc} = 300$, the same equation... (4) gives,

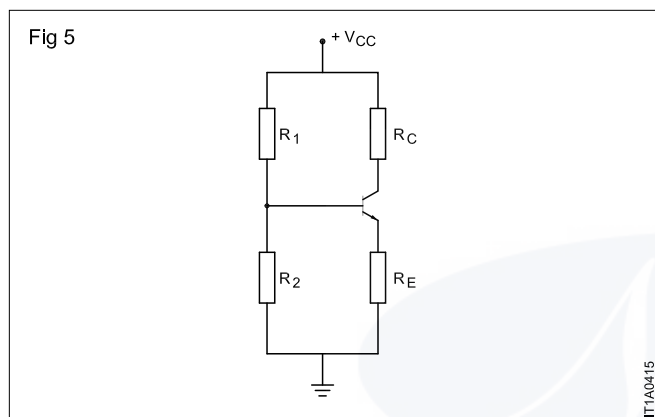
$$I_C = \frac{12 \text{ V} - 0.7 \text{ V}}{120 \Omega + 330 \text{ K}\Omega / 300} = 9.262 \text{ mA.}$$

Fig 4 summarizes the calculations by showing the dc load line and the two Q points. As can be seen, a 3:1 change in β_{dc} produces almost a 3:1 change in the collector current. This change is unacceptable as a stable-biased state.



TIP: For linear operation of the transistor, the base resistor R_B should be greater than $\beta_{dc} R_C$. A base resistance of less than $\beta_{dc} R_C$ produces saturation in an emitter-feedback-biased circuit.

3 VOLTAGE-DIVIDER bias: Fig 5 shows a typical voltage-divider bias. This type of biasing is also called the *universal bias* because, this is the most widely used type of biasing in linear circuits.



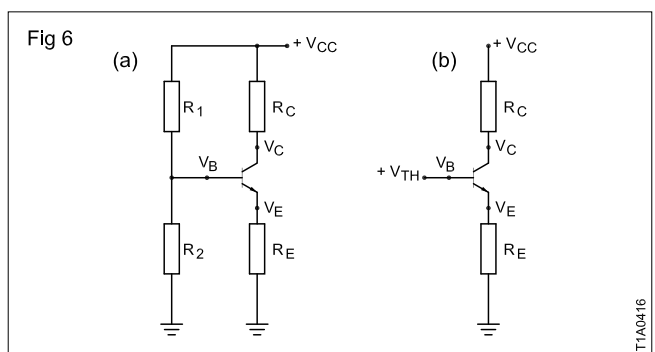
This type of biasing is known as *voltage divider bias* because of the voltage divider formed by resistors R_1 and R_2 . The voltage drop across R_2 should be such that it forward biases the emitter diode.

Emitter current in voltage divider bias

Assume that the base lead is open as shown in Fig 6b. Looking back at the unloaded voltage divider,

$$V_{TH} = \frac{R_2}{R_1 + R_2} V_{CC}$$

V_{TH} is known as the Thevinin's voltage. Refer reference books for Thevinin's theorem.



Now assume that, the base lead is connected back to the voltage divider as in Fig 6a. Then, voltage V_{TH} drives the base of the transistor. In other words, the circuit simplifies to Fig 6a and the transistor acts like the controlled current source.

Because the emitter is boot-strapped to the base,

$$I_E = \frac{V_{TH} - V_{BE}}{R_E}$$

The collector current I_C will be approximately equal to I_E .

Notice that β_{dc} does not appear in the formula for emitter current. This means that the circuit is not dependent on variations in β_{dc} . This means that the divider-biased transistor has a stable Q point.

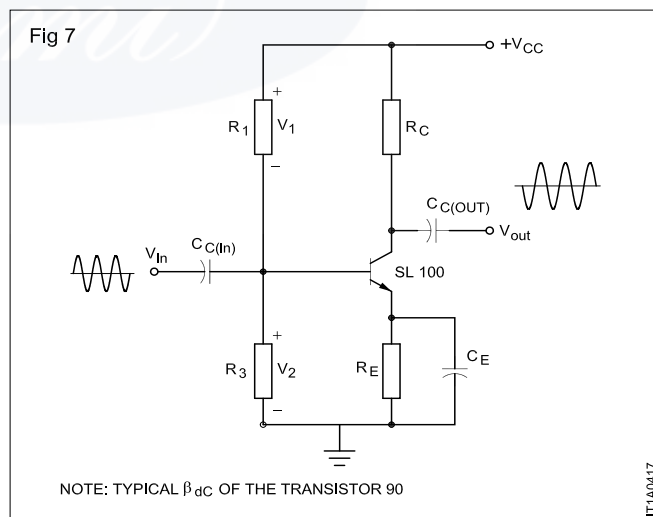
Because of the stable Q point, voltage-divider bias is the most preferred form of bias in linear transistor circuits. Hence, divider bias is used almost universally.

Simple voltage divider bias design guidelines

Fig 7 shows an amplifier using voltage-divider bias.

The capacitors C_C couple the ac signal into and out of the amplifier. The capacitor C_E is used to by-pass ac signals. A small ac input voltage drives the base, and an amplified ac output voltage appears at the collector.

In Fig 7, as far as the dc voltages and current is concerned, the capacitors appear like open circuits. Hence they can be neglected while designing the circuit for a stable Q point.



Design Step 1

Choose V_E equal to one-tenth of the chosen V_{CC} .

In Fig 7,

$$V_E = 0.1 V_{CC} \\ = 0.1 \times 12 \text{ V} = 1.2 \text{ V.}$$

This design rule is suitable for most circuits, but remember that this is only a guideline. It is not necessary that everyone uses this rule. So, do not be surprised to find emitter voltages at values different from one-tenth of the supply voltage.

Design Step 2

Fix the required value of I_C almost equal to I_E . Ensure that the worst case chosen I_C is less than $I_{C(max)}$ of the transistor given in the data book.

For Fig 7, fix $I_E \gg I_C = 10 \text{ mA}$.

$$\text{Therefore, } R_E = \frac{V_E}{I_E} = \frac{1.2 \text{ V}}{10 \text{ mA}} = 120 \Omega$$

Design Step 3

To locate the Q point at approximately the middle of the dc load line, fix $V_{CE} = 0.5 V_{CC}$.

$$\text{Therefore, } V_{CE} = 0.5 V_{CC} = 0.5 \times 12 \text{ V} = 6 \text{ V}.$$

$$\text{Hence, } V_{RC} = V_{CC} - V_{CE} - V_E = 12 - 6 - 1.2 = 4.8 \text{ V}.$$

$$\text{Therefore, } R_C = \frac{V_{RC}}{I_C} = \frac{4.8 \text{ V}}{10 \text{ mA}} = 480 \Omega$$

Choose the nearest 470W as R_C .

TIP : With the above design rule, without calculation, you can simply choose the value of R_C as,

$$R_C = 4 \times R_E = 4 \times 120 = 480 \Omega \text{ which is same as calculated in step 3.}$$

Design Step 4

To make a stiff voltage divider, apply the 10:1 rule,

$$\text{i.e., } R_2 \leq 0.1 b_{dc} R_E.$$

Take the typical value of b_{dc} from data book.

In Fig 7, with $b_{dc} = 90$

$$R_2 \leq 0.1 \times 90 \times 120 \Omega = 1080 \Omega.$$

Therefore, choose $R_2 = 1000 \Omega = 1 \text{ K}\Omega$

Design Step 5

Find the voltage divider voltages V_1 & V_2 .

$$V_2 = V_{BE} + V_E$$

$$\text{In Fig 7, } V_2 = 0.7 \text{ V} + 1.2 \text{ V} = 1.9 \text{ V}$$

$$\text{Therefore, } V_1 = V_{CC} - V_2.$$

$$\text{In Fig 7, } V_1 = V_{CC} - V_2 = 12 \text{ V} - 1.9 \text{ V} = 10.1 \text{ V}.$$

Calculate R_1 using the formula,

$$R_1 = \frac{V_1}{V_2} R_2.$$

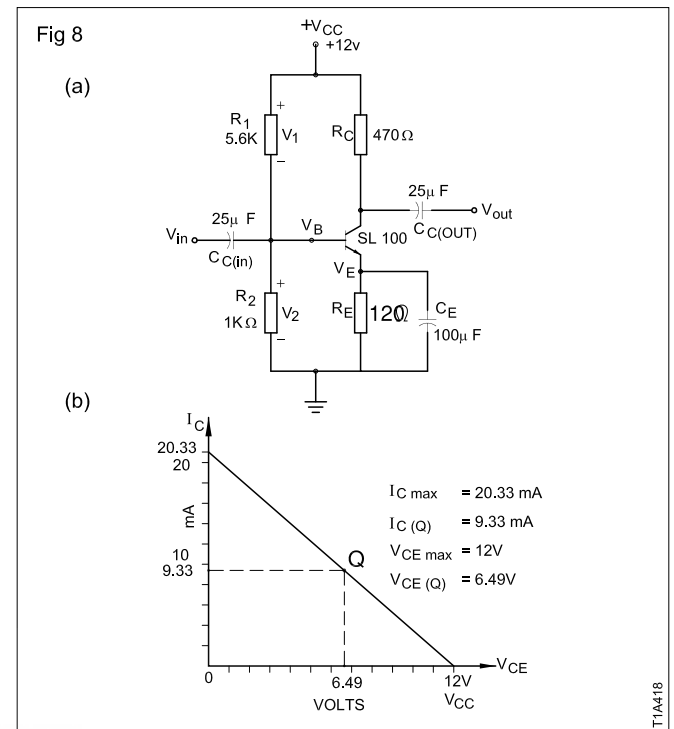
In Fig 7,

$$R_1 = \frac{10.1 \text{ V}}{1.9 \text{ V}} 1000 \Omega = 5.3 \text{ K}\Omega \approx 5.6 \text{ K}\Omega$$

This completes the design of the voltage-divider bias for the transistor amplifier at Fig 7.

Fig 8a shows the transistor amplifier with the designed values of components.

Fig 8b shows the dc load line and the Q point. As can be seen in Fig 8b, the Q point is near the middle of the dc load line. Hence, the designed circuit works in the linear portion of the transistor characteristic curve.



To cross check the above design we can calculate values of voltages and currents using the formulas given below;

$$V_E = \frac{R_2}{R_1 + R_2} \cdot V_{CC}$$

Let us call the parallel combination of R_1 and R_2 as R_{BB} .

$$I_B = \frac{V_E - V_{BE}}{R_{BB} + (\beta + 1) R_E} \quad (\text{See note given below})$$

$$I_C = \beta I_B$$

$$V_{CE} = V_{CC} - I_C (R_C + R_E)$$

The formula given for I_B is,

$$I_B = \frac{V_E - V_{BE}}{R_{BB} + (\beta + 1) R_E}$$

This formula has arrived as given below;

$$I_E = I_C + I_B$$

$$I_E = \beta I_B + I_B$$

$$I_E = (\beta + 1) I_B$$

In a divider bias, the effective resistance of $R_1 \parallel R_2$, denoted as R_{BB} is given by,

$$R_{BB} = \frac{R_1 \cdot R_2}{R_1 + R_2}$$

The effective bias voltage is represented as V_B is given by,

$$V_B = \frac{R_2}{R_1 + R_2} \cdot V_{CC}$$

Writing the Kirchhoff's voltage equation for the base emitter loop at Fig 8,

$$V_{CC} - I_B \cdot R_{BB} - V_{BE} - I_E R_E = 0 \quad \dots(1)$$

Note that R_{BB} is the effective resistance of parallel combination of R_1 and R_2 .

Since $I_E = (\beta + 1) I_B$

Rewriting equation...(1)

$$V_{CC} - V_{BE} - I_B \cdot R_{BB} - (\beta + 1) I_B R_E = 0$$

Solving the equation, we get, $I_B = \frac{V_{CC} - V_{BE}}{R_{BB} + (\beta + 1) R_E}$

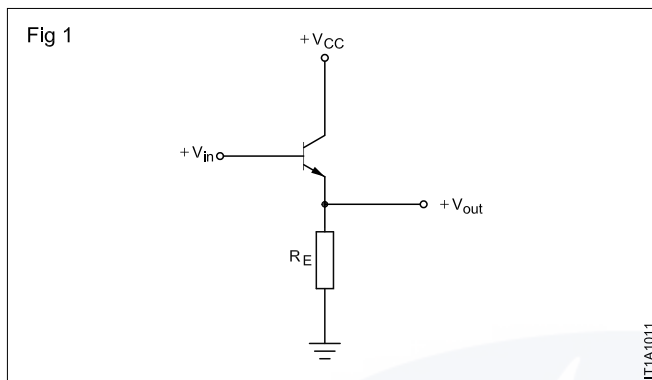


Emitter follower / Common Collector

Objectives: At the end of this lesson you shall be able to

- state the need for impedance matching
- state the popular application of an emitter follower amplifier configuration
- calculate the voltage gain, current gain, input impedance and output impedance of a transistor amplifier using circuit component values.

Fig 1 shows another important transistor amplifier configuration. In this configuration, unlike in a common emitter amplifier where the output is taken from the collector, the output is taken from the emitter terminal of the transistor as shown in Fig 1.



Emitter follower can be used, to match a high impedance source to a low impedance output load. Hence, the emitter follower configuration is frequently used as an impedance matching circuit than as an amplifier.

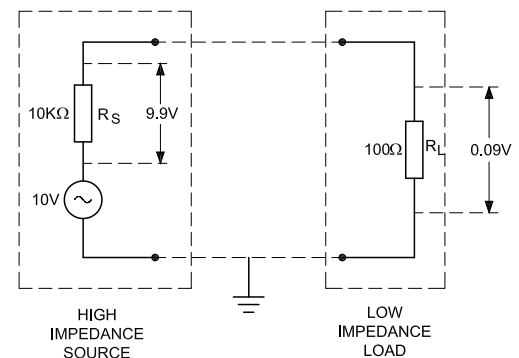
Need for impedance matching

When a high impedance source is connected to a low impedance load, then most of the ac signal of the source gets dropped across the internal impedance of the source itself resulting in a very small portion of the signal appearing across the required load as shown in Fig 2a. One way to overcome this problem, i.e. to have almost all the signal from the source to be developed across the load, is to use an impedance matching device or a circuit between the high impedance source and the low impedance load as shown in Fig 2b.

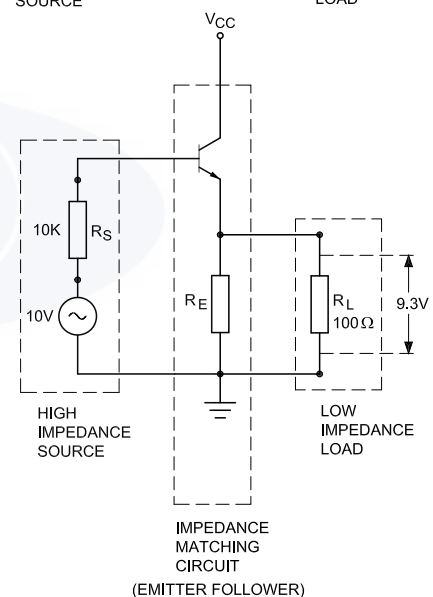
The circuit used for impedance matching in Fig 2b, is an emitter follower transistor amplifier. This is because, the emitter follower has a very high input impedance and a very low output impedance. This can be compared to that of a matching transformer where a load is matched to the source impedance for maximum power transfer.

An emitter follower circuit is also called a *common collector amplifier* because, the collector behaves as the common terminal for AC signal between the input and the output.

Fig 2
(a)



(b)



Voltage gain of an emitter follower

As can be seen in Fig 3, the DC output voltage of the emitter follower is $V_{out} = V_{in} - V_{BE}$

Since, V_{BE} is almost a constant value (0.7 for silicon, 0.3 for germanium) the emitter voltage follows the base voltage. It is because the emitter voltage follows the base voltage, this circuit is called **emitter follower**.

In Fig 3, if V_{in} is 3V, then $V_{out} = 2.3V$. If V_{in} is made 4V then V_{out} increases to 3.3V. This means that changes in V_{out} is in phase with changes in V_{in} . Therefore in an emitter follower the input and output signals are in phase as shown in Fig 3b. (recall, in a CE amplifier the input and output are 180° out of phase.)

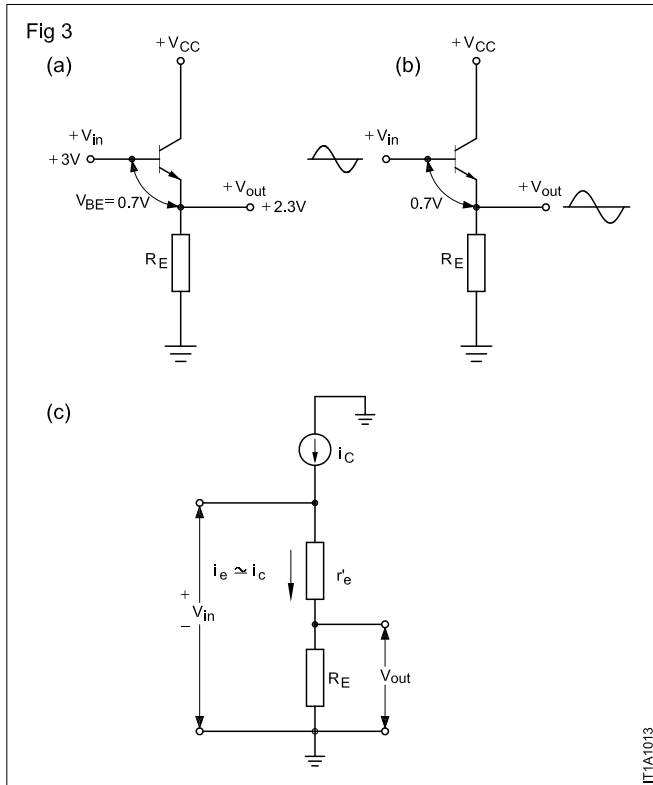


Fig 3c shows the AC equivalent circuit of the emitter follower shown in Fig 3a. The ac output voltage V_{out} is given by,

$$V_{out} = i_e R_E$$

Since the ac input voltage V_{in} is given by, $V_{in} = i_e (R_E + r'_e)$ the voltage gain A_v of the emitter follower is,

$$A_v = \frac{V_{out}}{V_{in}} = \frac{i_e R_E}{i_e (R_E + r'_e)} = \frac{R_E}{R_E + r'_e} \quad \dots\dots\{1\}$$

In equation {1}, since the denominator will always be higher than the numerator, the value of voltage gain A_v will always be less than 1.

But since the value of r'_e is very small compared to R_E , the value of A_v approaches unity. We can therefore say that the voltage gain of the emitter follower is unity.

In Fig 3a, if $R_E = 4.7KW$ and $r'_e = 25W$ then,

$$A_v = \frac{R_E}{R_E + r'_e} = \frac{4700}{4700 + 25} = 0.995 \approx 1$$

Input impedance of emitter follower

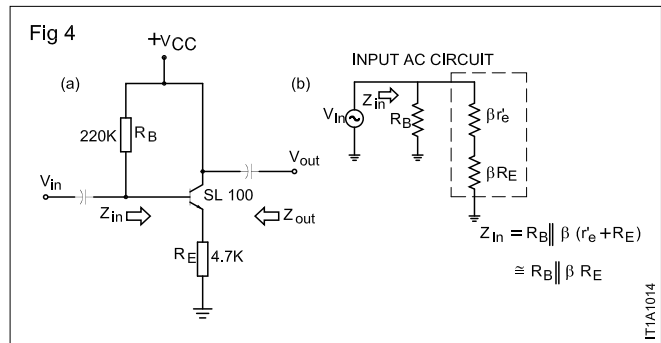
The input impedance of the emitter follower shown in Fig 3 is given by,

$$Z_{in} = b(R_E + r'_e) \quad \dots\dots\{2\}$$

Since r'_e will be generally very small compared to R_E , equation ...{2} can be simplified as,

$$Z_{in} = b R_E$$

Fig 4a shows a practical emitter follower circuit using fixed biasing. The total input impedance including the biasing resistor R_B in parallel with the input impedance can be found as follows;



Writing the AC equivalent of the input of the emitter follower shown in Fig 4b, the input impedance Z_{in} is given by,

$$Z_{in} = R_B \parallel b(r'_e + R_E) \quad \dots\dots\{3\}$$

If r'_e is neglected, then, $Z_{in} = R_B \parallel bR_E$

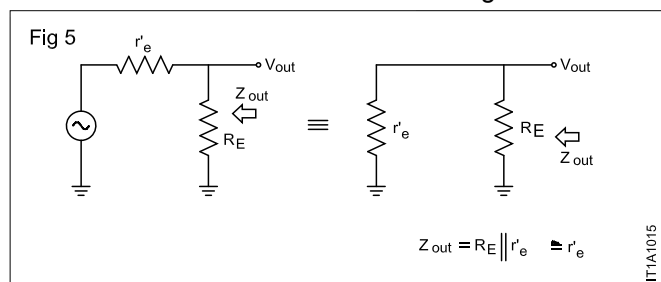
Equation 3 indicates that the input impedance of a typical emitter follower is decided by the DC biasing resistance R_B . Hence, while designing an emitter follower to match a high source impedance, the values of R_B should be suitably chosen.

Example: In the emitter follower at Fig 4, if b of transistor is 100, $R_B = 220 k$ and $R_E = 4.7k$ the input impedance will be,

$$\begin{aligned} Z_{in} &= R_{in} = R_B \parallel bR_E \\ &= 220 K \parallel b R_E \\ &= 220 K \parallel (100 \times 4.7K) \\ &= 149.85 K \Omega \approx 150 K \Omega \end{aligned}$$

Output impedance of emitter follower

Fig 5 shows an AC equivalent or AC model of the output section of an emitter follower shown in Fig 4a.



Analysing Fig 5, the output impedance Z_{out} is given by,

$$Z_{out} = R_E \parallel r'_e \quad \dots\dots\{4\}$$

Since R_E is usually a large resistance compared to r'_e , R_E in equation (4) can be neglected. Therefore, the output impedance of an emitter follower is approximately,

$$Z_{out} = r'_e$$

Example 1: Find the output impedance of the emitter follower shown in Fig 4 assuming $r'_e = 33W$,

$$Z_{out} = r'_e \parallel R_E \approx r'_e = 33W.$$

Current gain in emitter follower

Although the voltage gain A_V of emitter follower is approximately unity, the current gain of emitter follower is high and is given by the equation;

$$A_I = \frac{\beta R_E}{(R_B + \beta R_E)} \quad \dots\dots\{5\}$$

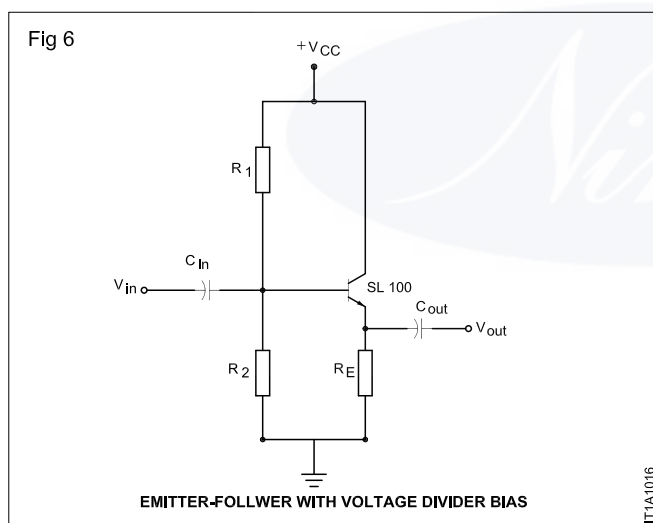
Example 2: In the emitter follower shown in Fig 4, if β of the transistor is 100, then the current gain of the emitter follower is given by,

$$A_I = \frac{\beta R_E}{(R_B + \beta R_E)} \\ = \frac{(100)(220K)}{(220K) + (100)(4.7K)} = 31.88.$$

The current gain of the emitter follower can also be found as follows;

$$A_I = A_V \frac{Z_{in}}{R_E} = (0.995) \frac{150K}{4.7K} = 31.72.$$

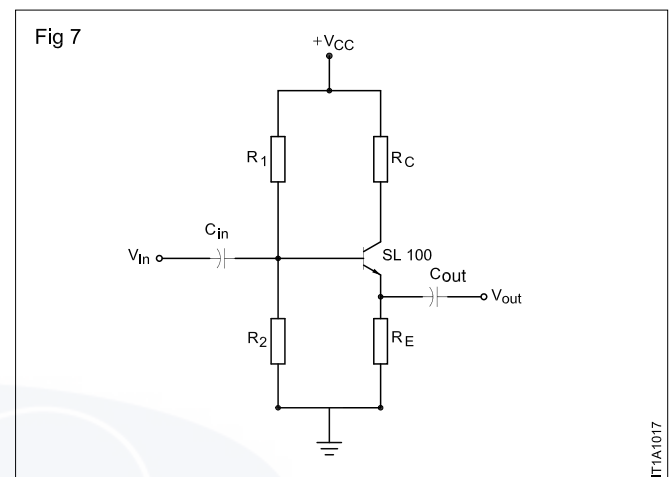
Fixed bias was used in the emitter follower shown in Fig 4. Any other dc biasing could also be employed such as voltage divider bias as shown in Fig 6.



When voltage divider bias is used in an emitter follower, the equations for finding A_V , Z_{in} , Z_{out} and A_I remain the same, except for, that the fixed biased resistor R_B replaced by $R_1 \parallel R_2$.

Fig 7 which includes a collector resistor R_C also provides the input-output characteristics of an emitter follower. The input impedance Z_i and the output impedance Z_o are not effected by R_C since it is not reflected into the base or emitter equivalent network of the circuit. The only use of R_C will be to fix the quiescent collector current or Q point of the transistor.

Appendix-D gives a summary of the important parameters of common collector or emitter follower amplifier.

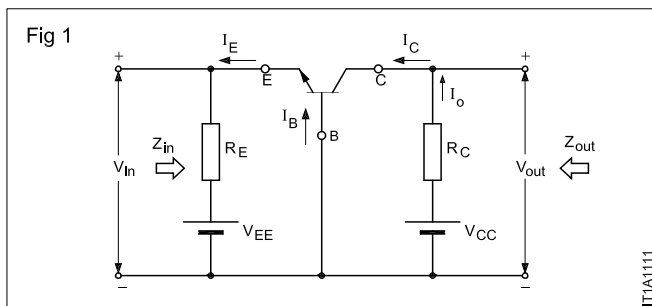


Common base amplifier

Objectives: At the end of this lesson you shall be able to

- calculate the voltage gain, current gain, input impedance and output impedance of a common base amplifier
- list the typical applications of common base amplifiers. using circuit component values.

Fig 1 shows the typical circuit schematic of a **common base amplifier** (CB-amplifier).

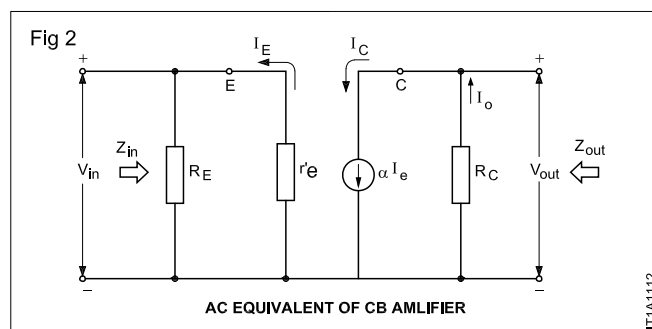


Common base amplifiers have a current gain of less than 1. Recall, the emitter current and the collector current of a transistor are almost equal. In a common base amplifier since the input current is I_E and the output current is I_C , the current gain, symbolically represented as α (alpha), is given by,

$$\text{Current gain } (\alpha) = \frac{\text{Output current}}{\text{Input current}} = \frac{I_C}{I_E}$$

Since $I_E > I_C$, α will always be less than 1.

Fig 2 shows the AC equivalent of a common base amplifier, shown in Fig 1.



From Fig 2, the input impedance Z_{in} of CB amplifier is given by,

$$Z_{in} = R_E \parallel r'_e \quad \dots\dots\{1\}.$$

Since R_E is generally much greater than r'_e , eqn..1 can be simplified as,

$$Z_{in} = r'_e$$

The equation for Z_{in} indicates that, the input impedance of a CB amplifier is very low and almost equal to the AC resistance r'_e of the emitter diode (recall r'_e will be generally 25Ω).

Referring to the AC equivalent network of the CB amplifier, the output impedance Z_{out} of CB amplifier is given by,

$$Z_{out} = R_C \quad \dots\dots\{2\}.$$

Equation 2 indicates that the output impedance of a CB amplifier is relatively high, of the order of kilo ohms (because you can fix the value of R_C as you wish!).

From Fig 2, the output voltage V_{out} is

$$V_{out} = I_o R_C = I_C R_C$$

$$\text{Since } \alpha = \frac{I_C}{I_E}, I_C = \alpha \cdot I_E$$

$$\text{Therefore, } V_{out} = I_C R_C = \alpha \cdot I_E R_C \quad \dots\dots\{3\}.$$

Since, $I_E = \frac{V_{in}}{r'_e}$ equation 3 can be written as,

$$V_{out} = \alpha \left(\frac{V_{in}}{r'_e} \right) R_C$$

The voltage gain A_v of CB amplifier is given by,

$$A_v = \frac{V_{out}}{V_{in}} = \alpha \frac{V_{in}}{r'_e} R_C \frac{1}{V_{in}} = \alpha \frac{R_C}{r'_e} \quad \dots\dots\{4\}$$

Since r'_e is very small compared to R_C , the voltage gain A_v of the CB amplifier is quite high.

The power gain A_p of the CB amplifier is given by,

$$A_p = A_i A_v$$

Power gain A_p will be medium because although A_i is less than or equal to 1, A_v of the CB amplifier is quite high.

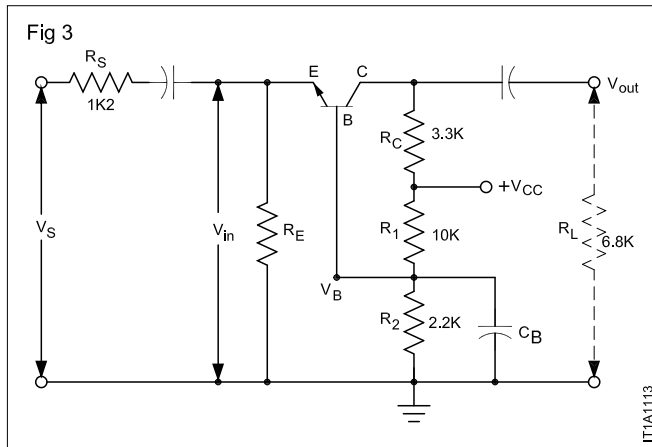
Input/Output phase relationship

The input and output of a common base amplifier are in phase with each other. This can be found experimentally.

Fig 3 shows a CB amplifier with voltage divider bias.

In Fig 3, the base of the transistor is at AC ground due to the bypass capacitor C_B . The input signal drives the emitter and the output is taken from the collector. The biasing resistors R_1 , R_2 will have negligible effect on the input impedance. Therefore, the input impedance of the CB amplifier is approximately equal to r'_e itself.

The input V_{in} to the amplifier is given by (note that C_B bypass R_2 for AC signal),



The voltage at the base (at T_1) is given by,

$$\begin{aligned}
 V_B &= \frac{R_2}{R_1 + R_2} \cdot V_{CC} \\
 &= \frac{2.2K}{10K + 2.2K} \cdot 12V \\
 &= 2.16 \text{ volts.}
 \end{aligned}$$

The emitter current I_E is given by,

$$\begin{aligned}
 I_E &= \frac{V_B - V_{BE}}{R_E} \\
 &= \frac{2.16 - 0.7}{1.2K\Omega} \\
 &= 1.22 \text{ mA.}
 \end{aligned}$$

Therefore, r'_e is given by,

$$\begin{aligned}
 r'_e &= \frac{25mV}{I_E} \\
 &= \frac{25mV}{1.22 \text{ mA}} \\
 &= 20.5 \Omega
 \end{aligned}$$

Input impedance Z_{in} is given by,

$$Z_{in} \triangleq r'_e = 20.5 \Omega$$

The voltage gain A_v is given by,

$$\begin{aligned}
 A_v &= \frac{R_C}{r'_e} \\
 &= \frac{3.3K}{20.5} \\
 &= 160.97 \triangleq 161
 \end{aligned}$$

The output impedance Z_{out} is given by,

$$Z_{out} \triangleq R_C = 3.3K\Omega$$

$$\begin{aligned}
 V_{in} &= \frac{r'_e}{R_S + r'_e} V_s \\
 &= \frac{20.5 \Omega}{1K\Omega + 20.5 \Omega} 500mV \\
 &= 10 \text{ mV.}
 \end{aligned}$$

Therefore, the unloaded output voltage V_{out} is given by,

$$\begin{aligned}
 V_{out(\text{No load})} &= A_v \cdot V_{in} \\
 &= 161 \times 10 \text{ mV} \\
 &= 1610 \text{ mV} \\
 &= 1.61 \text{ volts.}
 \end{aligned}$$

The output voltage of the amplifier with load R_L is given by,

$$\begin{aligned}
 V_{out(\text{Load})} &= \frac{R_L}{R_C + R_L} \times V_{out(\text{No load})} \\
 &= \frac{6.8K}{3.3K + 6.8K} \times 1.61V \\
 &= 1.08V
 \end{aligned}$$

Class Room Assignment: Calculate the output voltage of the CB amplifier (as done in step above) if load resistor R_L was,

- (i) $R_L = 3.3K$
- (ii) $R_L = 10K$ and,
- (iii) $R_L = 100K$

Appendix-D gives a summary of the important parameters of common base transistor amplifiers.

Field effect transistor and its types

Objectives: At the end of this lesson you shall be able to

- explain the difference between bi-polar transistors and Field effect transistors
- write the basic construction and symbol used
- explain the theory of operation of FETs
- explain a typical FET a.c voltage amplifiers.

FIELD EFFECT TRANSISTORS (FET)

The main difference between a Bi-polar transistor and a FET is that,

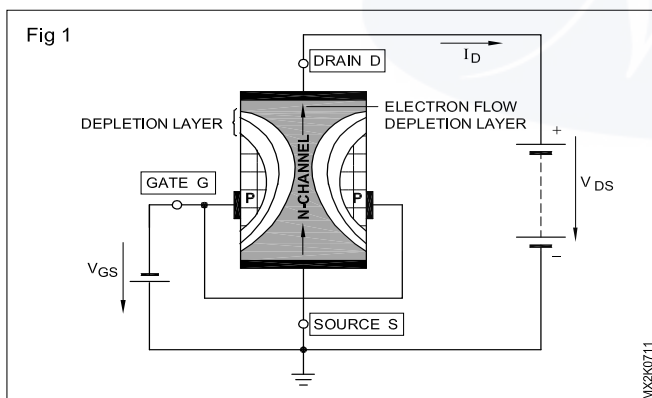
Bi-polar transistor is a current controlled device

In simple terms, it means that the main current in a bi-polar transistor (collector current) is controlled by the base current.

FET is a voltage controlled device

This means that the voltage at the gate(similar to base of a bi-polar transistor) controls the main current.

In addition to the above, in a bi-polar transistor (NPN or PNP), the main current always flows through N-doped and P-doped semiconductor materials. Whereas, in a FET the main current flows either only through the N-doped semiconductor or only through the P-doped semiconductor as shown in Fig 1.



If the main current flow is only through the N-doped material, then such a FET is referred as a N-channel or N-type FET. The current through the N-doped material in the N-type FET is only by electrons.

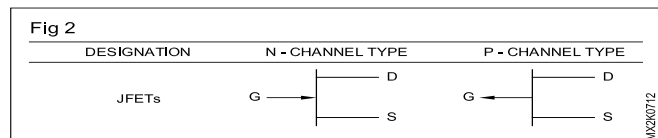
If the main current flow is only through the P-doped material, then such a FET is referred as a P-channel or P-type FET. The current through the P-doped material in the P-type FET is only by Holes.

Unlike in bipolar transistors in which the main current is both by electrons and holes, in contrast in FETs depending on the type(P or N type) the main current is either by electrons or by holes and never both. For this reason FETs are also known as Unipolar transistors or Unipolar device.

There are a wide variety of FETs. In this lesson MOSFET (Metal Oxide Semiconductor FET) and one of the fundamental type called as Junction Field Effect Transistor (JFET) are discussed.

Junction Field effect Transistor(JFET)

It is a three terminal device and looks similar to a bi-polar transistor. The standard circuit symbols of N-channel and P-channel type FETs are shown in Fig 2.



Construction

As shown in Fig 3a, a n-Channel JFET has a narrow bar of n-type. To this, two p-type junctions are diffused on opposite sides of its middle part Fig 3a. These diffused junctions form two P-N diodes or gates. The n-type semiconductor area between these junctions/gates is called channel. The diffused P regions on opposite sides of the channel are internally connected and a single lead is brought out which is called gate lead or terminal. Direct electrical connections are made at the two ends of the bar. One of which is called source terminal, S and the other drain terminal, D.

A p-channel FET will be very similar to the n-channel FET in construction except that it uses P-type bar and two N type junctions as shown in Fig 3b.

FET notation listed below are essential and worth memorizing,

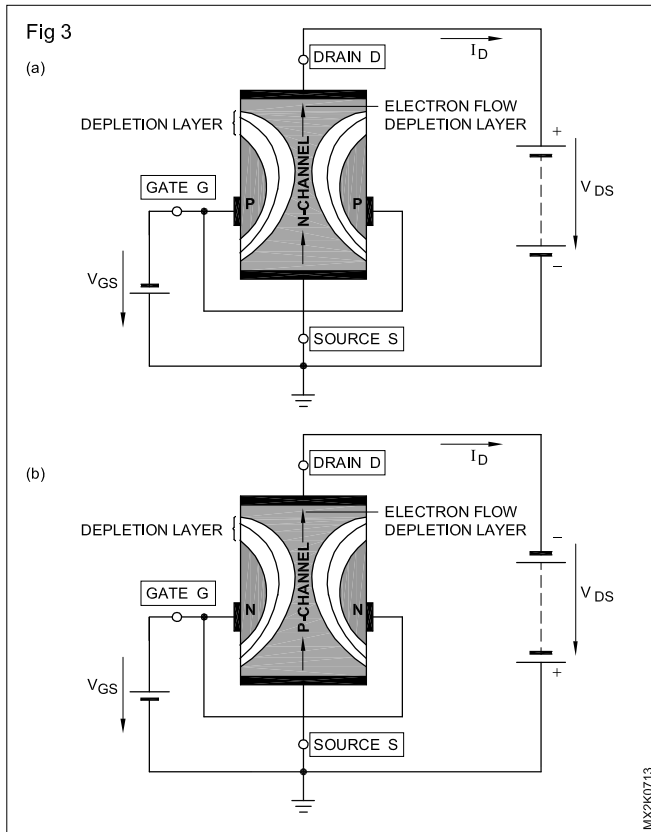
- 1 Source terminal: It is the terminal through which majority carriers enter the bar(N or P bar depending upon the type of FET).
- 2 Drain terminal: It is the terminal through which majority carriers come out of the bar.
- 3 Gate terminal: These are two internally connected heavily doped regions which form two P-N junctions.
- 4 Channel: It is the space between the two gates through which majority carriers pass from source to drain when FET is working(on).

Working of FET

Similar to Bipolar transistors, the working point of adjustment and stabilization are also required for FETs.

Biassing a JFET

- Gates are always reverse biased. Therefore the gate current I_g is practically zero.



- The source terminal is always connected to that end of the supply which provides the necessary charge carriers. For instance, in a N-channel JFET source terminal S is connected to the negative of the d.c power supply. And, the positive of the d.c power supply is connected to the drain terminal of the JFET.

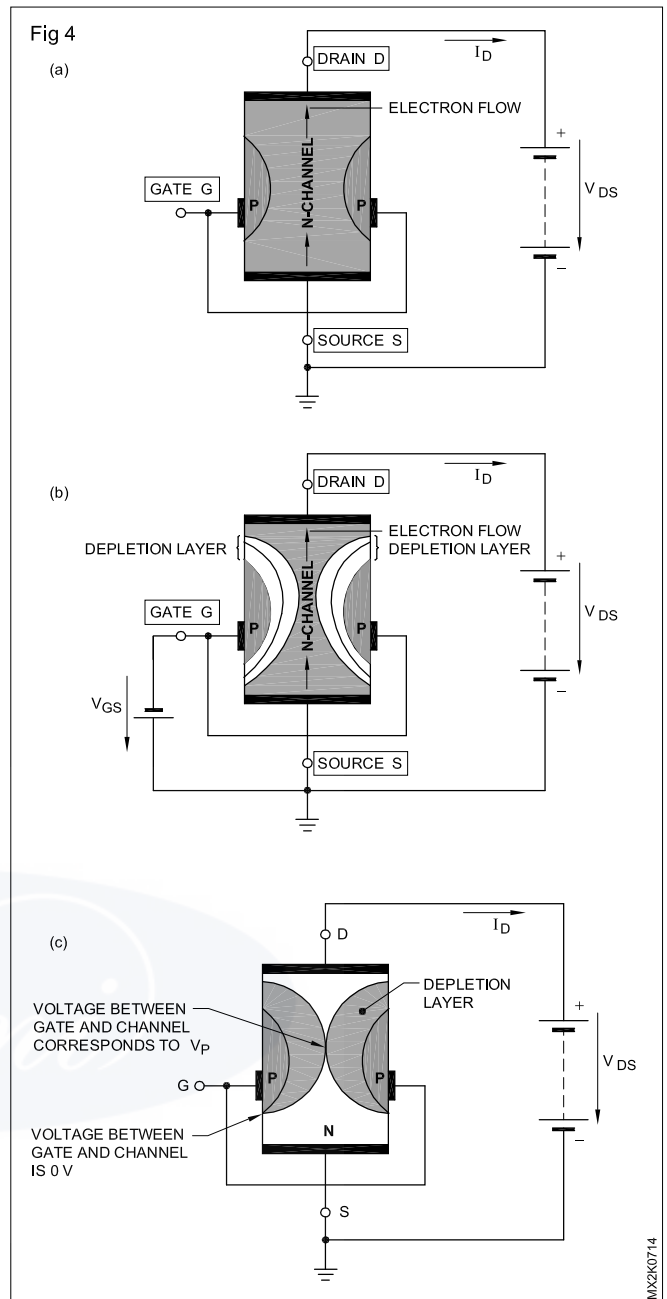
where as in a P channel JFET, Source is connected to the positive end of the power supply and the drain is connected to the negative end of the power supply for the drain to get the holes from the P-channel where the holes are the charge carriers.

Let us now consider an N channel JFET, the drain is made positive with respect to source by voltage V_{DS} as shown in Fig 4a. When gate to source voltage V_{GS} is zero, there is no control voltage and maximum electron current flows from source(S) - through the channel - to the drain(D). This electron current from source to drain is referred to as Drain current, I_D .

When gate is reverse biased with a negative voltage(V_{GS} negative) as shown in Fig 4b, the static field established at the gate causes depletion region to occur in the channel as shown in Fig 4b.

This depletion region decreases the width of the channel causing the drain current to decrease.

If V_{GS} is made more and more negative, the channel width decreases further resulting in further decrease in drain current. When the negative gate voltage is sufficiently high, the two depletion layers meet and block the channel cutting off the flow of drain current as shown in Fig 4c. This voltage at which this effect occurs is referred to as the Pinch off voltage, V_p .



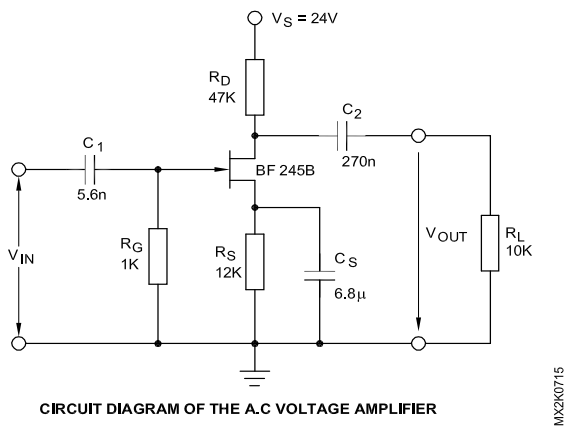
Thus, by varying the reverse bias voltage between gate and source($-V_{GS}$), the drain current can be varied between maximum current (with $-V_{GS}=0$) and zero current(with $-V_{GS}=\text{pinch off voltage}$). So, JFET can be referred as a voltage controlled devices.

P channel JFET operates in the same way as explained above except that bias voltages are reversed and the majority carrier of channel are holes.

FET ac voltage amplifier

Fig 5 shows FET amplifier circuit. The amplification factor of the amplifier can be varied by varying the drain and source resistance. Pots can be connected in series for this purpose.

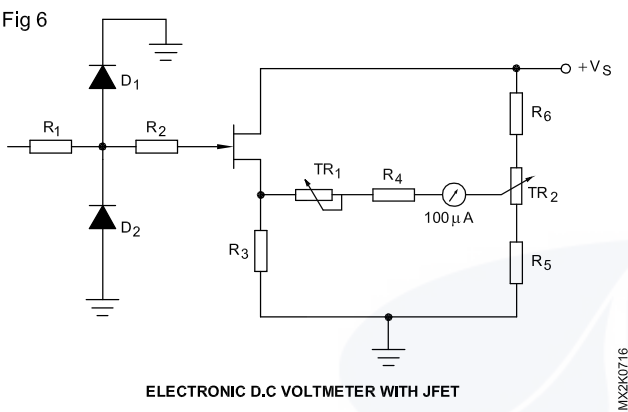
Fig 5



Electronic d.c. voltmeter using JFET.

By virtue of their very high input resistance, field-effect transistors can also be used satisfactorily as d.c. voltage amplifiers in electronic voltmeters. This is shown in Fig 6.

Fig 6



In Fig 6, the JFET is operated with a common drain. With its source resistance, the JFET forms one branch of a bridge circuit. The 100-μA measuring mechanism is connected in the bridge diagonal. Zeroing is carried out using trimmer Tr2, with the input short-circuited by setting the bridge current to zero.

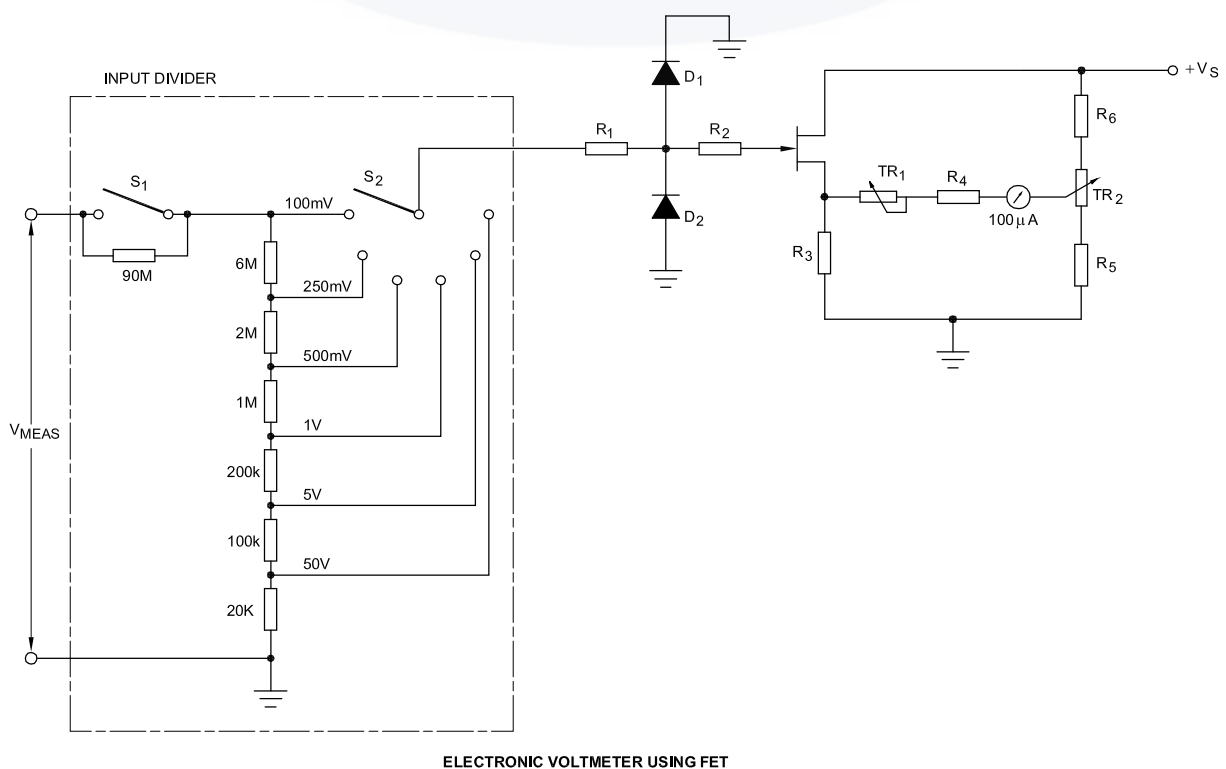
Full-scale deflection of the 100-μA instrument takes place at an input voltage of 100mV. Balancing to full scale deflection is carried out using trimmer Tr1. Diodes D1 and D2 act as overload protection. They become conducting when a voltage in excess of 0.7V occurs after input resistance R1. The gate bias is generated by a resistor R2.

If a d.c. voltage measuring signal is delivered via input resistor R1 to the gate of the JFET, its drain current increases. Its channel resistance R_{DS} is thus reduced and the bridge ceases to be in equilibrium. A current flows through the instrument, whose deflection is proportional to the input measuring voltage.

The electronic voltmeter described above can be extended into a multi-range instrument by connecting a multi-stage voltage-divider in front of it as shown in Fig 7.

The above design of the voltage divider produces measuring ranges with full-scale deflections for $U_{in} = 100\text{mV}$ to $U_{in} = 50\text{V}$. These measuring ranges can be selected using switch S2. By opening switch S1, an additional resistance of 90 MW is inserted between the input socket and the voltage divider. This increases all the measuring ranges by a factor of 10.

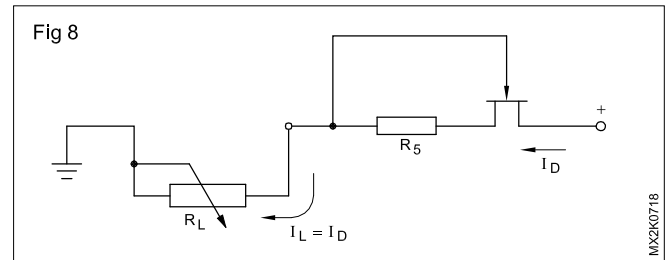
Fig 7



When switch S1 is closed, the electronic voltmeter thus has six measuring ranges with full-scale deflections of 100mV, 250mV, 500mV, 1V, 5V and 50V. The input resistance is in this case 10 MW. When switch S1 is open, six measuring ranges are obtained, with full-scale deflections of 1V, 2.5V, 5V, 10V, 50V and 500V. The input resistance is then 100MΩ.

3 Constant current source

Constant-current sources are circuits with a very high internal resistance, which supply almost constant currents over a wide voltage range. Constant-current sources can be very easily constructed using field-effect transistors. Use is made of the high output resistance R_{DS} of FETs when operated in the pinch-off region. The drain-source voltage U_{DS} has only a very small influence on the magnitude of the drain current. Fig 8 shows a constant-current source with a JFET.



When using constant-current sources with JFETs, it is absolutely essential to ensure that the limiting values U_{DSmax} , I_{Dmax} and P_{tot} cannot be exceeded.

Constant-current sources with FETs have the great advantage that additional auxiliary voltages are not essential and that they can be treated as two-terminal networks. They can therefore be inserted in a circuit like any normal resistor.

Types of FET(s)

Objectives: At the end of this lesson you shall be able to

- discuss different types field effect transistors
- discuss special type of MOSFETs.

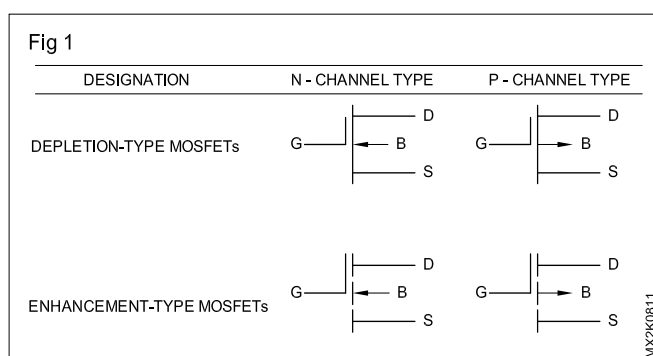
INTRODUCTION TO DIFFERENT TYPES OF FETs

There is a wide variety and types of FET. Basically, There are two groups:

- 1 Junction FETs (discussed in previous lesson)
- 2 MOSFETs (MOS = Metal-Oxide-Semiconductor FET).

JFETs as already discussed are further subdivided into n-channel types and p-channel types. In the case of MOSFETs, on the other hand, we must differentiate between enhancement-type and depletion-type MOSFETs, for which again n-channel types and p-channel types occur. There are also a few special types.

The circuit symbols of MOSFET's are shown in Fig 1.



One characteristic property of all FETs is their high input resistance. In the case of JFETs it is around $10^9 \Omega$, and in the case of MOSFETs roughly $10^{15} \Omega$.

In addition to the standard types of JFET and MOSFET, there are a few special types of further developments of MOSFETs. These are the dual-gate MOSFET, the VMOSFET and SIPMOSFET.

Dual-gate MOSFETs are used to construct amplifiers, whose amplification factors can be controlled within wide limits. VMOSFETs and SIPMOSFETs, on the other hand, were developed specially for amplifying and switching high powers of the order of 5KW. Their great advantage is the very low control power required.

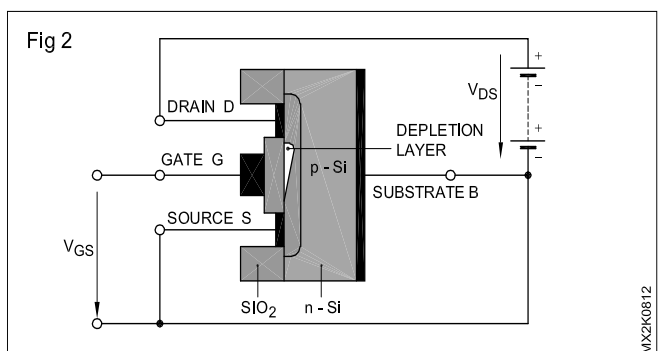
Metal-Oxide-Semiconductor Field Effect Transistors (MOSFETs)

In MOSFETs, control is via an insulating layer instead of a junction (as in JFETs). This insulating layer is generally made of silicon dioxide, from which the very name MOSFET is derived (Metal Oxide Semiconductor). Some times the MOSFETs are also referred to as Insulated-gate FET, for which the abbreviation used are IFET or IGFET.

Depletion-type MOSFETs

Construction and mode of operation

Fig 2 shows the construction of a depletion MOSFET of the n-channel type.



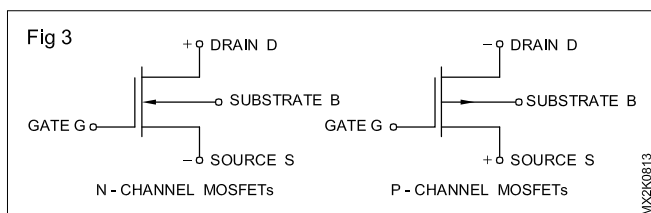
Here, two highly doped n-zones are diffused into a p-doped silicon plate, which is referred to as the substrate, and are provided with junction-free drain and source connections. Between the two zones there is a thin n-doped channel, which produces an electrical connection between the source and drain without an external field-action. This channel is covered by an insulating layer of silicon dioxide (SiO_2), to which a metal electrode is applied as the gate connection.

If a voltage U_{DS} is applied between source and drain, at $U_{GS} = 0V$ an electron current flows from the source electrode via the n-channel to the drain electrode. If, however, a negative voltage is applied to control electrode G, the electrons present in the n-channel are forced out of the vicinity of the gate electrode, so that a zone depleted of charge carriers is produced there. This causes a constriction of the n-channel and consequently also a reduction of its conductivity. If the gate voltage becomes more negative, the conductivity of the channel is reduced, as is consequently also the drain current I_D . Another peculiarity of depletion-type MOSFETs is that they can also be controlled with a positive gate-voltage. Charge carriers are then drawn out of the p-doped substrate into the n-channel and its conductivity is increased even further, compared with the conductivity at $U_{GS} = 0V$.

Designations and circuit symbols

The same designations are used for the connections of MOSFETs as they are for JFETs, i.e. source, drain and gate. MOSFETs, however, have another electrode, which is referred to as the substrate connection. Together with the semiconductor material of the channel, this substrate forms a p-n junction, which can be used as a second control-electrode. It is then led out of the casing, like the other electrodes. In a number of versions, however, the substrate electrode is connected directly to the source connection in the casing, which rules out the additional control possibility.

Fig 3 shows the circuit symbols for depletion-type n-channel MOSFETs and p-channel MOSFETs. For the n-channel type, the arrow points towards the line representing the channel; in the case of the p-channel type, on the other hand, it points away from the line representing the channel. The continuous line representing the channel indicates that it is a depletion-type MOSFET.

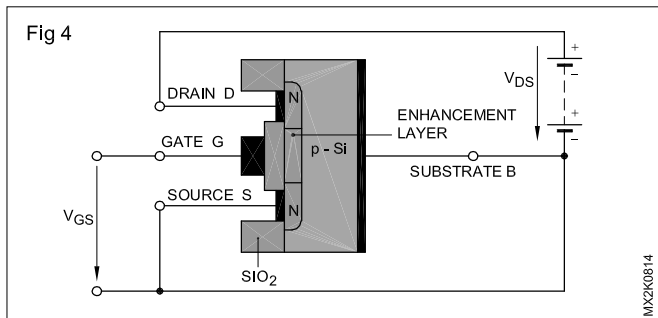


N-channel MOSFETs are operated with a positive drain-source voltage. They have a considerably greater practical significance than p-channel MOSFETs, which require a negative drain-source voltage for their operation.

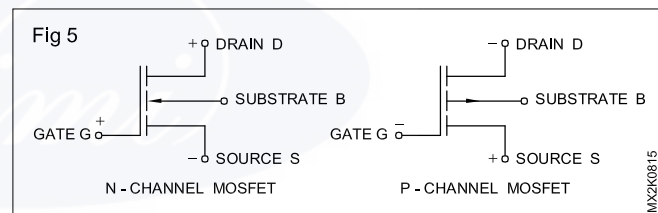
Enhancement-type MOSFETs

Construction and mode of operation

Enhancement-type MOSFETs have a similar technological construction to the depletion types. Without the external action of a field, however, no conducting channel exists between the drain connection and the source connection, so that at $U_{GS} = 0V$, no drain current can flow. Fig 4 shows the construction of an enhancement-type n-channel MOSFET.



The same circuit designations are used for the four electrodes of the enhancement-type MOSFET as they are for the depletion types: drain, source, gate and substrate. The circuit symbols used are different. The line representing the channel in the circuit symbol is discontinuous for an enhancement-type MOSFET. This indicates that no drain current I_D flows at $U_{GS} = 0V$. The circuit symbols for the two types of enhancement MOSFET are given in Fig 5.

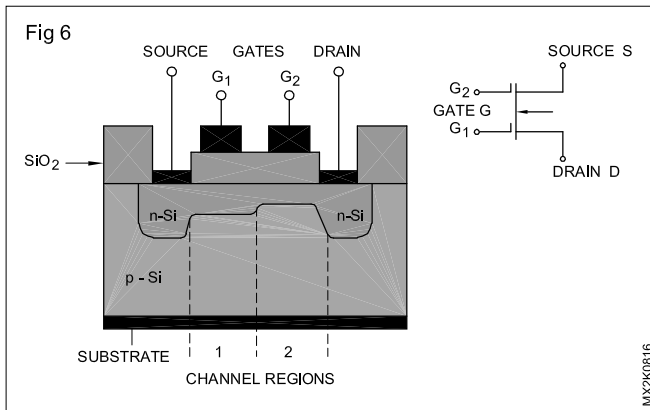


Enhancement-type MOSFETs are only rarely produced as individual transistors. Their construction and working principle are, however, widely used in integrated MOS switching circuits.

SPECIAL TYPES OF MOSFET

Dual-gate MOSFET

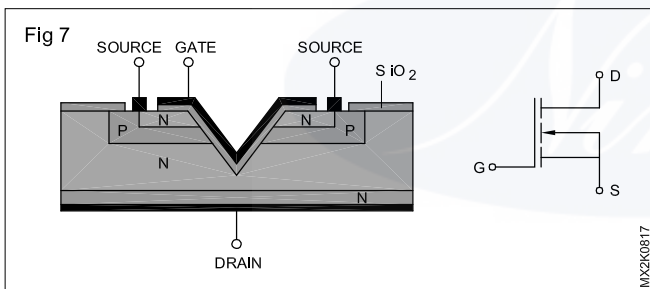
The dual-gate MOSFET is a special type of depletion MOSFET. It has two series-connected channel regions as the current path. The conductivity of each of these two channel regions can be independently controlled via its own gate. The construction and circuit symbol of a depletion-type dual-gate MOSFET of the n-channel type are reproduced in Fig 6. Because of the four connections, this special type is also referred to as a "MOSFET tetrode".



VMOSFET

The field-effect transistors dealt with so far, can handle only relatively small powers during amplification or switching. The reason for this is the relatively long channel of approximately 5 μm with a forward resistance of approximately $1\text{K}\Omega$ to $10\text{K}\Omega$. With present-day manufacturing techniques, it is possible to produce a vertical structure for field-effect transistors, instead of the customary horizontal sequence of layers. Consequently, higher allowable currents and voltages are obtained, so that considerably greater powers can be amplified or switched.

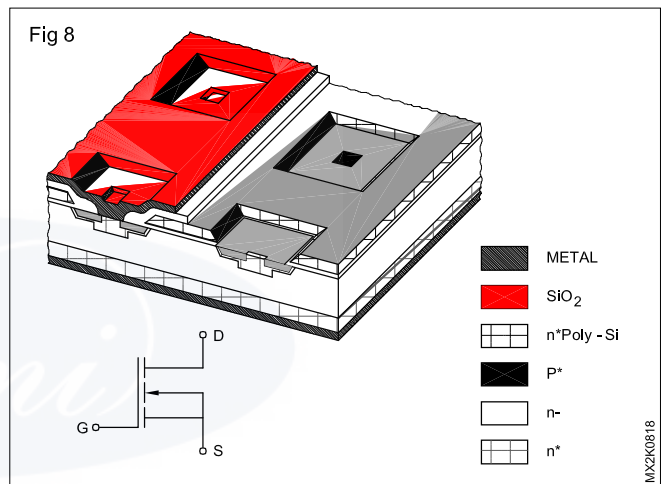
Fig 7 shows the construction of an enhancement-type n-channel VMOSFET and the associated circuit symbol.



SIPMOSFET

The name SIPMOSFET is an abbreviation for the designation Siemens Power-MOSFET. Here again, these are MOSFETs for high powers. These are achieved by a rather different layer structure from that in VMOSFETs. Fig 8 shows the layer construction and circuit diagram of a SIPMOSFET.

As shown in Fig 8, the SIPMOSFET is constructed by planar structure technology, with a horizontal layer sequence. There is a weakly conducting n-layer on a low-resistance n⁺-crystal. Cell shaped n⁺-conducting source areas are arranged in p-conducting recesses at its surface. A gate electrode, embedded and insulated in quartz, covers the surface between the source cells. These individual source-cells are connected in parallel by metallizing plating. Such a SIPMOSFET is thereby produced from the parallel connection of several thousand individual MOS elements. A high permissible power dissipation is thereby achieved. SIPMOSFETs are enhancement-type field-effect transistors.



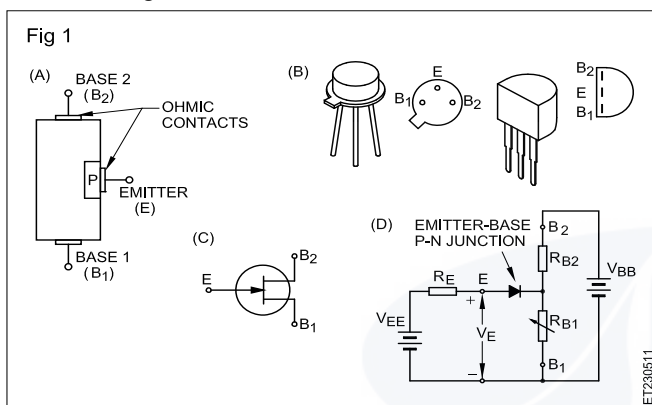
Unijunction Transistor (UJT) its applications

Objectives: At the end of this lesson you shall be able to

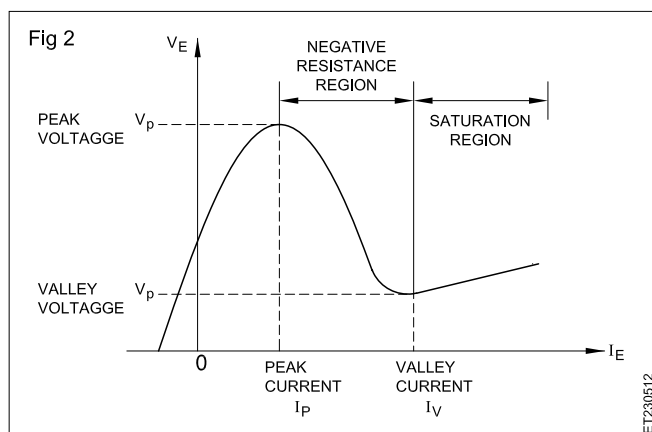
- aware of the construction and working principle of UJT
- able to make a quick test of UJT
- able to list the important specification of UJT
- list and explain the application of UJTs.

Unijunction transistor(UJT) is a three terminal semiconductor device as shown in fig 1a. In its appearance it looks like a transistor as shown in Fig 1b. As shown in Fig 1a, it consists of two layers(a P-layer and a N-layer) and therefore it has only one junction(hence its name, unijunction).

The symbol of UJT and its electrical equivalent circuit is shown in Fig 1c and 1d.



UJT is a special semiconductor device because it exhibits negative resistance characteristics as shown in Fig 2. The details of the characteristics is discussed in subsequent paragraphs.



Construction of UJT

From its construction a UJT can be described as a double based diode. UJT is fabricated on a N-type silicon bar with ohmic contact for the two base terminals as shown in Fig 1a. The P-type emitter is heavily doped and is on one side of the N-type bar. An ordinary ohmic, non rectifying contact is made at the two ends of the N-type bar. These contacts are called Base-1 and Base-2.

When packaged UJT looks like a small signal Bipolar junction transistor with power ratings less than 1 Watt. For example the 2N 2646 and 2N 2647 UJT's are available in the modified TO-18 case style as shown in Fig 1b.

Equivalent circuit of UJT

The electrical equivalent circuit of UJT is shown in fig 1d. The resistance between the B_1 and B_2 terminals is called the inter-base resistance R_{BB} . The N-type silicon bar serves as a resistance divided into two parts R_{B1} and R_{B2} by the PN junction. The total of the internal R_{B1} and R_{B2} is the interbase resistance R_{BB} . Value of R_{BB} is typically in the range of 4 to 10 Kohms. Also r_{B1} usually a little greater than r_{B2} because the emitter is a little closer to B_2 .

The interbase resistance R_{BB} is measured with the emitter open.

$$R_{BB} = R_{B1} + R_{B2} \text{ at } I_E = 0.$$

Operation of UJT

The DC supply polarities for a UJT to function is shown in Fig 3. As can be seen from fig 3, B_2 is connected to +ve and B_1 to ground. As a result current (conventional) flows from B_2 to B_1 . This conduction results in a voltage gradient along the N-type silicon bar. Therefore there is a voltage in the region of the emitter junction (V_E) which is positive with respect to ground. The magnitude of this voltage is given by the simple voltage divider action between R_{B1} and R_{B2} .

$$V_E \text{ or } (V_{RB1}) = \frac{R_{B1}}{R_{B1} + R_{B2}} V_{BB} = \eta V_{BB} \quad \dots[1]$$

The Greek letter η (eta) is called the intrinsic stand-off ratio. This is an important data of any UJT and is invariably mentioned in all UJT data sheets. From the above equation, intrinsic stand-off ratio η (eta) is given by,

$$\eta = \frac{R_{B1}}{R_{B1} + R_{B2}} \quad \dots[2]$$

Typical values of η vary from about 0.5 to 0.8 for most UJT's.

The value of η is important because, knowing η and the applied voltage across the base terminals V_{BB} , voltage across R_{B1} can be calculated using equation [1].

For example, if $h = 0.65$ and $V_{BB} = 12$ Volts,

then, $V_{RB1} = h V_{BB} = 0.65 \times 12 \text{ V} = 7.8$ volts.

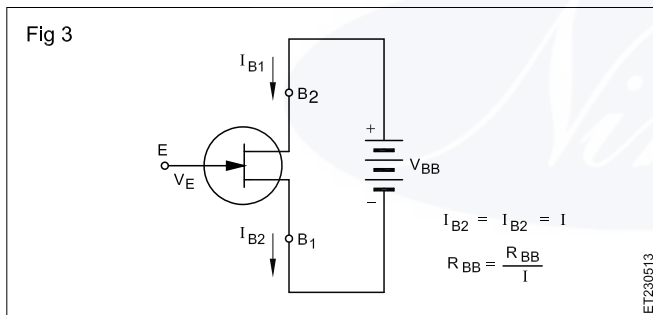
Voltage V_{RB1} represents the reverse bias voltage at diode D. In order for an emitter current I_E to flow, the emitter voltage V_{EB1} should be above V_{RB1} by about 0.7 volts (internal barrier potential for a silicon diode). This emitter voltage V_{EB1} that will cause the diode to be forward biased and conduct emitter current is usually designated as V_p .

The value of V_p can be calculated using the formula,

$$V_p = h V_{BB} + 0.7 \text{ volts}$$

In the example considered above, $V_p = 7.8 \text{ V} + 0.7 \text{ V} = 8.5$ Volts.

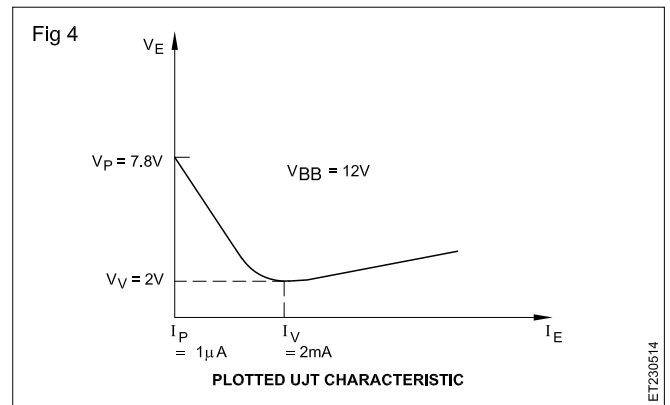
When V_p is raised to 8.5 volts, diode D conducts. This means I_E flows through the R_E , diode D, R_{B1} to ground. This sudden rush of current through the highly doped N-type R_{B1} reduces the resistance of R_{B1} (remember R_{B1} is not a fixed value resistor, it is the resistance of the N-type semiconductor). Because of the reduced R_{B1} (from a few thousand ohms to a value as low as 40 to 50 ohms) results in a small voltage drop across R_{B1} . Hence emitter voltage V_{EB1} drastically decreases to around 1 volt (from 8.5V before the diode conduction to around 1V when the diode starts conducting). This drop in the value of V_{EB1} is shown in Fig 3. This point is sometimes referred to as the on-state condition.



Since an increase in the emitter current results in decrease in emitter voltage (V_{EB1}), this region of the UJT characteristics is referred to as the negative resistance region.

After a certain point referred to as the valley point, any further increase in the emitter biasing voltage V_{EE} /emitter current I_E will start increasing the voltage drop across the emitter and base-1. This region is called as the UJT's saturation region. This will occur when the rate of hole injection is so great as to build up a positive space charge in the region (for further details refer reference books listed). It should be noted that after the valley point, the emitter current should not be increase beyond the rated maximum forward emitter current value (typically around 50mA) beyond which the diode D may break down.

A comparison of the UJT theoretical characteristics (as in Fig 2) and actual characteristics which can be plotted is shown in Fig 4.



Some of the important terms associated with UJT characteristics are given below;

- Peak Point Current (I_p) - minimum amount of emitter current to place the UJT in negative resistance region.
- Valley Current (I_v) - maximum allowable emitter current within the negative resistance region.
- Valley voltage (V_v) - minimum voltage that can maintain the UJT in its negative resistance region.

Quick test of UJT using OHM meter

From the construction of a UJT it can be seen that, Emitter and Base-1 with Base-2 open, behaves as a PN diode. Therefore, when tested using a ohm meter this should show low resistance when forward biased and high resistance when reverse biased.

Similarly Emitter and Base-2 with Base-1 open behaves as a PN diode and hence the same forward and reverse bias test using a ohm meter can be carried out to confirm its good condition.

To carry out a quick test of a given UJT, check forward and reverse bias conditions of the two diodes of UJT as given in above two paragraphs.

Typical UJT Specifications

UJT specification as can be seen in any data manual is given below. 2N 2646 UJT is taken only as a sample for understanding the specifications. Specifications for other UJT will almost be in the same format. However the manufacturers data sheets give more details than what is listed below;

Type	Device	I_P	I_V	R_{BBO}	Eta(h)	Package
2N 2646	UJT-P	5 μ A	4mA	15Kohms	0.60	see page X

Application of UJTs

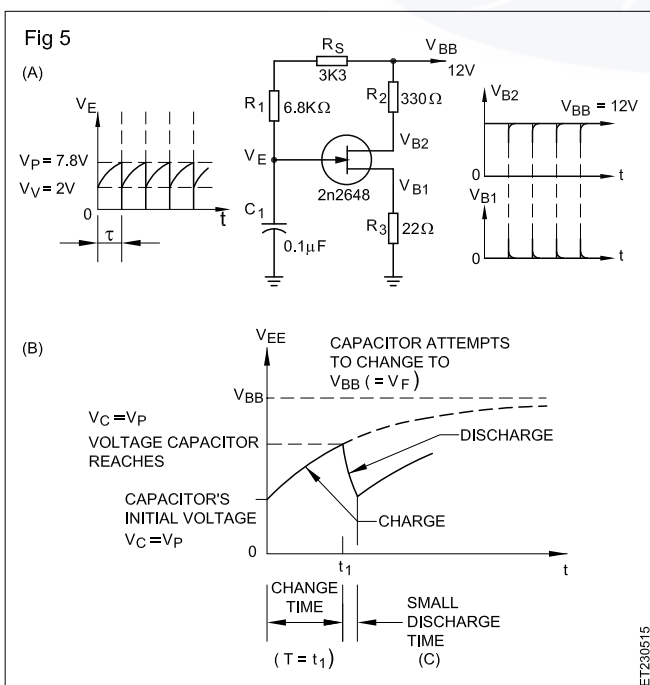
UJTs are employed in a wide variety of circuits involving electronic switching and voltage or current sensing applications. These include,

- Triggers for thyristers
- As oscillators
- As pulse and Saw tooth generators
- Timing circuits
- Regulated Power supplies
- Bistable circuits and so..on.

The most common and popular application of UJT is the Relaxation oscillator. Fig 5a shows a practical relaxation oscillator using 2N 2646 UJT.

As the voltage V_{BB} is supplied, capacitor C is charged via Resistors R_S and R_1 . If the voltage V_C across the capacitor cross the peak point voltage(V_P) of UJT, the UJT goes into conduction.

The sooner UJT goes into conduction, the charged capacitor C discharges rapidly as shown in Fig 5b via the low inner base resistance R_{B1} and R_3 . This conduction of UJT and the discharge of C through the emitter-Base1 of UJT results in a sudden rush of current through R_3 and hence the voltage across R_3 increases sharply as shown in Fig 5.



By discharging the voltage across it, the voltage across the capacitor becomes smaller than the valley point voltage V_V . Because of this, UJT cuts-off once again. Because the UJT is cut-off, there is no current through R_3 and hence the voltage across it (output voltage) becomes zero as shown in Fig 4.

Once the UJT is cut-off, capacitor C starts charging again through R_S and R_1 . When the charged voltage again crosses the value of V_P , UJT turns-on again and the cycle repeats resulting in continuous pulse wave form at the output(across R_3).

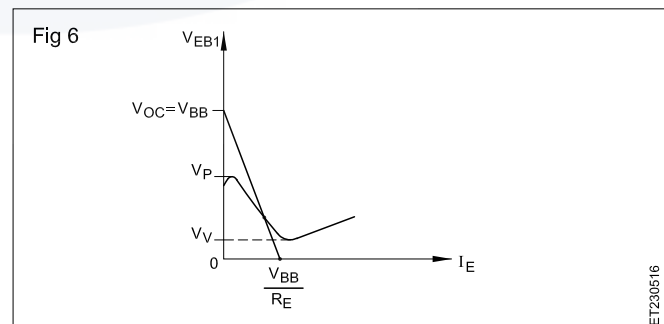
The frequency of oscillation of the UJT oscillator depends on,

- [1] Time constant $t_{ow}(t)$ given by $t_{ow}(t) = R_E \times C$
- and [2] Value of intrinsic standoff ration η of the UJT.

The frequency of oscillation(f_o) of the UJT relaxation oscillator is given by the formula,

$$f_o = \frac{1}{T} \approx \frac{1}{R_E C \ln \left[\frac{1}{(1-\eta)} \right]}$$

For a UJT to function properly in an oscillator, its DC load line must cross the negative resistance region of its emitter characteristics as shown in Fig 6.



An important thumb rule while designing an relaxation oscillator using UJT is given below;

Minimum and Maximum values of R_E to ensure oscillations,

$$R_{E(max)} = \frac{V_{BB} - V_P}{I_P}$$

$$R_{E(min)} = \frac{V_{BB} - V_V}{I_V}$$

Working of SCR, TRIAC and DIAC

Objectives: At the end of this lesson you shall be able to

- aware of construction and principle of working of SCR
- able to list the specifications and make a quick test of a given SCR
- different ways of triggering triac and quick testing of a triac
- use of triac for fullwave control of ac
- know the working and use of a DIAC
- know the method of quick testing a DIAC.

INTRODUCTION TO INDUSTRIAL ELECTRONICS

Industrial electronics is concerned primarily with electronics as applied to industries such as industrial equipments, controls and processes. An important application of electronics in industries is in controlling of machinery.

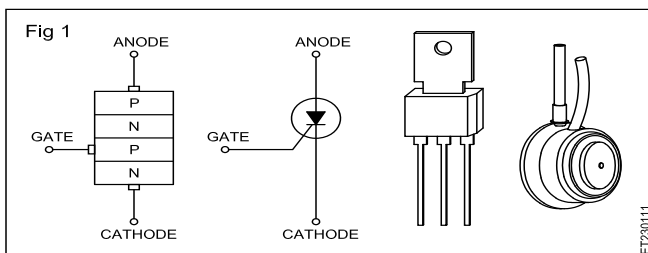
In communication electronics, domestic & entertainment electronics, generally, the electronic devices operate with currents in the order of microamperes to milliamperes. For industrial applications, most frequently, devices are required to handle currents in the range of few amperes to several hundreds of amperes. This, therefore calls for high power electronic devices. One such high power electronic device frequently used in industrial electronic application is the SCR. SCRs can be used to run, dc motors from an ac power source, control power tool speed, also to control motor speeds of small appliances like, mixers and food blenders, illumination control, temperature control and so on.

Silicon Controlled Rectifier(SCR)

Before Silicon controlled rectifiers were invented(1956), a glass tube device called Thyatron was used for high power applications. Silicon controlled rectifier (SCR) is the first device of the thyristor family. The term thyristor is coined from the expression Thyatron-transistor. SCR is a semiconductor device. SCR does the function of controlled rectification. Unlike a rectifier diode, SCR has an additional terminal called the gate which controls the rectification(gated silicon rectifier).

The basic principle application of SCRs is to control the amount of power delivered to a load(motor, lamp, etc.,).

A rectifier diode will have one PN junction. SCRs on the other hand will have two PN junctions (P-N-P-N layers). Fig 1 shows the electrical symbol, basic construction and a typical SCR packages.



Basic operation of SCR

When a gate current is applied to the gate terminal, forward current conduction commences in the SCR(latched into conduction). When the gate current is removed, the forward current through the SCR **does not cut-off**. This means, once the SCR is latched into conduction, the gate loses control over the conduction. The current through the SCR can be turned off only by reducing the current through it(load current) below a critical value called the **Holding current**.

Fig 2 shows how an SCR can be gated into conduction or turned off.

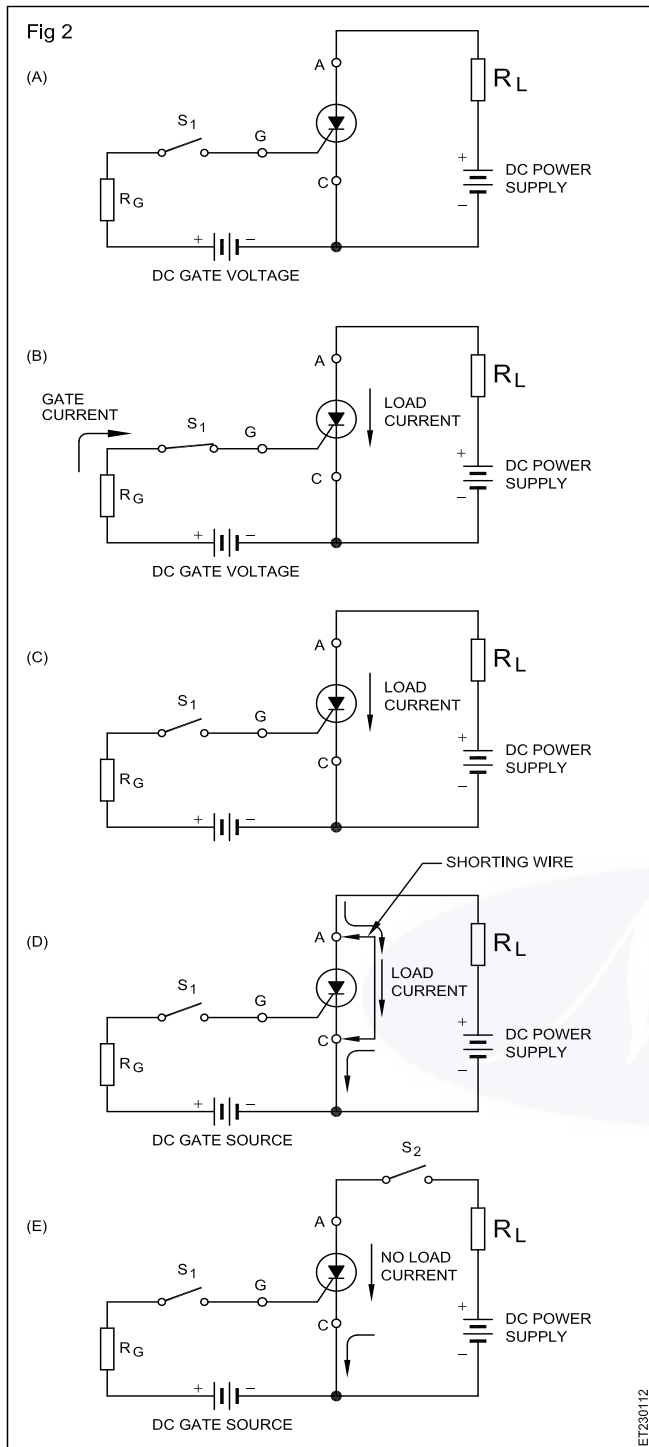
In Fig 2a, with switch S1 open the SCR is in OFF state and no current is flowing through the load.

In Fig 2b, when S1 is closed, a small gate current(around 1/1000 or less compared to load current) turns-ON (fires) the SCR. A heavy load current starts flowing through the SCR and load R_L .

In Fig 2c, when S1 is opened, gate current becomes zero. This will have no effect on the current through the SCR and the heavy load current continues to flow through the SCR.

In Fig 2d, if a shorting wire is placed across the anode and the cathode terminals, the current through the SCR gets by-passed and all the current starts flowing through the shorted wire instead through the SCR. This means the current through the SCR is reduced below the rated holding current(minimum current required through SCR to keep it latched). This turns-OFF the SCR. Even when the shorting wire is removed the SCR remains to be in OFF state.

Fig 2e shows an alternative method of turning-OFF the SCR. In this instead of shorting the anode and cathode terminals of the SCR, the load current is cut-off by opening the Switch S2. This reduced the current through the SCR below the holding current and thus turns- OFF the SCR. Once the SCR is turned_ OFF, the SCR does not turn-ON even if the switch S2 is closed. To make the SCR fire again, with the switch S2 closed, the gate current should be made to flow by closing the switch S1.



Since the SCR does not conduct in the reverse direction, the anode of the SCR should always be positive with respect to cathode for conduction.

Important features of SCR,

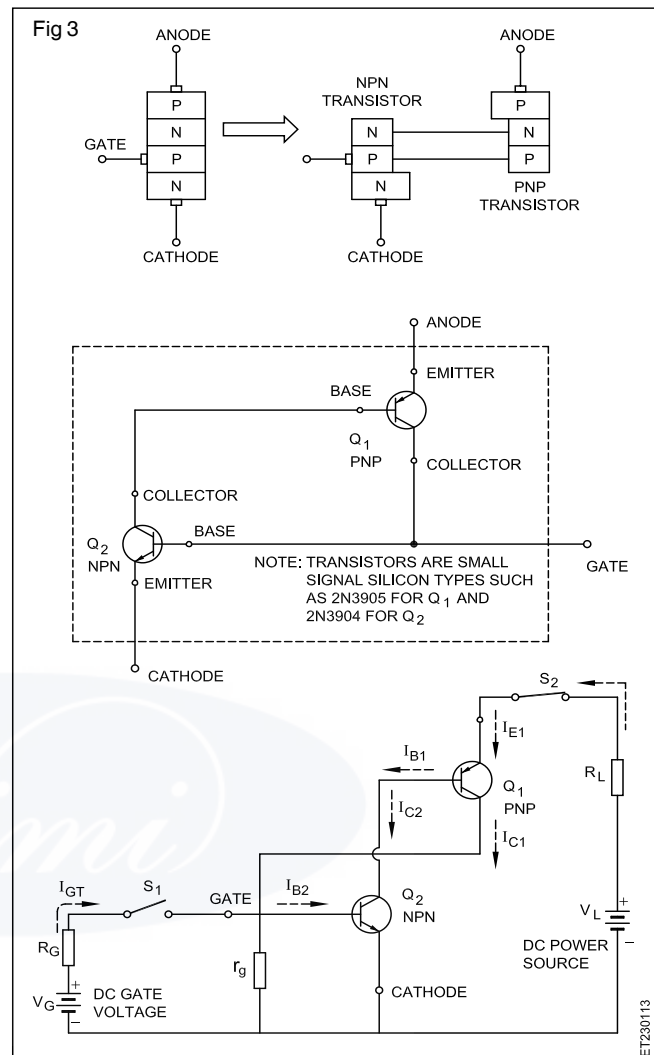
- A very small gate current will control the switching of a large load current.

Example: In 2N1597 SCR, a gate current of about 10mA will switch a load current of 1.6

- A gate current pulse of short duration (typically 1 to 2μs) is sufficient to turn on the forward load current of SCR.

Understanding SCR as two interconnected transistors

For a clear understanding of how a SCR works, the four layer PNPN construction of SCR can be visualized as two independent transistors interconnected as shown in Fig 3a.



In Fig 3c, in the initial state with S_1 open, both Q_1 and Q_2 are in cutoff and hence there is no current through the load R_L .

When S_1 is closed, the gate supply V_G causes a small base current (I_{B2}) to flow from the base of Q_2 . This small base current turns on Q_2 and fairly a large collector current I_{C2} flows. At the same time, the emitter-base junction of Q_1 is forward biased and turns on Q_1 . Since Q_1 collector current flows through the base of Q_2 , this causes I_{C2} to increase. This increase in the base and collector currents of Q_1 and Q_2 becomes regenerative and thus takes Q_1 and Q_2 to saturation. At this time even if the base externally applied gate trigger pulse is removed, since the collector current of each transistor supplies the required base current to each other transistor, both the transistors remain to be in saturation. In this condition, the SCR acts almost like a short circuit/closed switch with a resistance as small as 0.1 ohms or less.

To bring the SCR to OFF state, since the externally applied gate has no further control of the SCR, the only way to turn-off is to reduce the collector currents of Q1 and Q2 to such a level that they will not sustain the regenerative action. This can be done by breaking the load current by switching off S2 or by placing a short across the anode and cathode of the SCR as discussed earlier. By shorting or by opening S2, the current through the SCR (collector currents of the interconnect transistors) is reduced below the rated **holding current**.

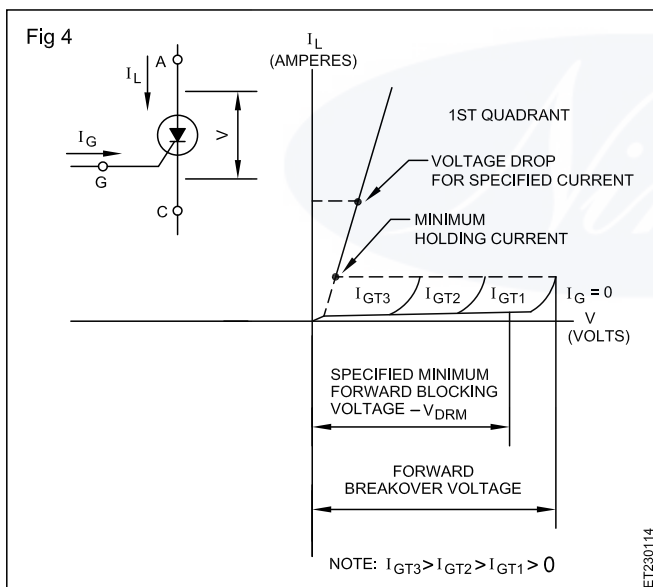
Holding current is that minimum value of current required to sustain the regenerative action in the interconnected transistors in the absence of the gate trigger.

The resistor r_g connected across the base and emitter of Q2 is to provide immunity from false gate triggering signals such as noise spikes etc. Recall that such a connection at the input circuit of a transistor reduces the input (gate) impedance of the transistor.

Typical VI Characteristics of SCR

SCR operation in first quadrant with DC supply

Typical Voltage-Current (VI) characteristics of a SCR is illustrated in Fig 4.



Considering that the SCR is operated with DC supply and the device is always forward biased (+ve to anode and -ve to cathode). The VI characteristics of a typical SCR will be as shown in Fig 4. This is called as the first quadrant operation of the SCR.

When SCR is used with ac supply, then the VI characteristics in the third quadrant will also be discussed.

Abnormal Operation of SCR

With Gate current $I_G = 0$,

- With a moderate Anode Cathode voltage, only a very small leakage current (can be neglected) flows through the SCR.

- If the anode cathode voltage is increased beyond a maximum forward blocking voltage (see fig 4), the SCR will break into forward condition and large current starts flowing through the SCR. This destroys the SCR breaking down the central pn junction. Therefore, the anode cathode voltage of SCR should always be less than the forward breaking voltage.

Break over voltage: It is that anode-cathode voltage at which the PN junctions of the SCR breaks down thus forcing the SCR into conduction.

Normal operation of SCR

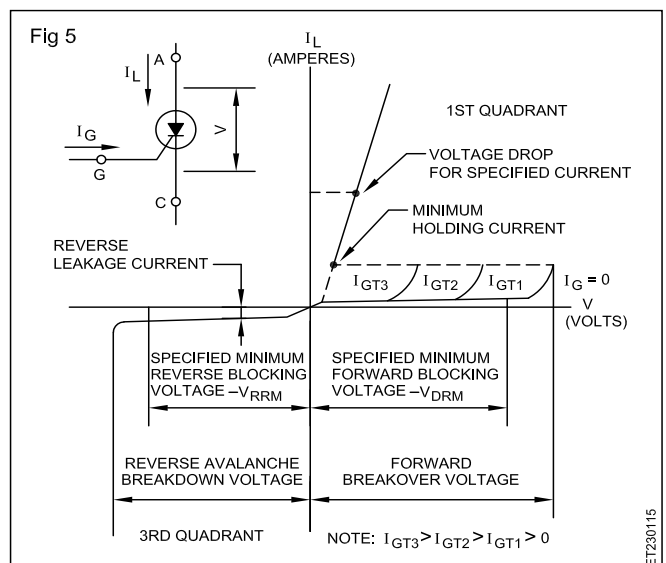
With small gate trigger currents I_{G1} or I_{G2} or I_{G3} (see Fig 4),

- With a small gate trigger current say I_{G1} , the SCR need a lower anode cathode voltage to fire the SCR on and a large current flows through the SCR. The amount of current through the SCR is limited only by the external load connected.
- With a slightly larger gate trigger current I_{G2} ($I_{G2} > I_{G1}$), the SCR gets fired even at a lesser anode cathode voltage.

Thus, a lower anode-cathode voltage is required to set the SCR into conduction with higher values of gate trigger current and vice-versa.

- Once the SCR is fired, it remains in conduction as long as the SCR current/load current is not allowed to drop below the "holding current value". As and when the SCR is required to stop conduction, the load current should be brought below the holding current value either by switching-off the load circuit (load current becomes zero) or by reversing the applied voltage polarity to the anode-cathode or by other suitable means.

Figure 5 illustrated the VI characteristics of SCR in first and third quadrant. The third quadrant operation basically illustrated the operation of the SCR when in reversed biased condition. The operation of the SCR in First and Third quadrant occurs when the SCR is operated using AC supply.

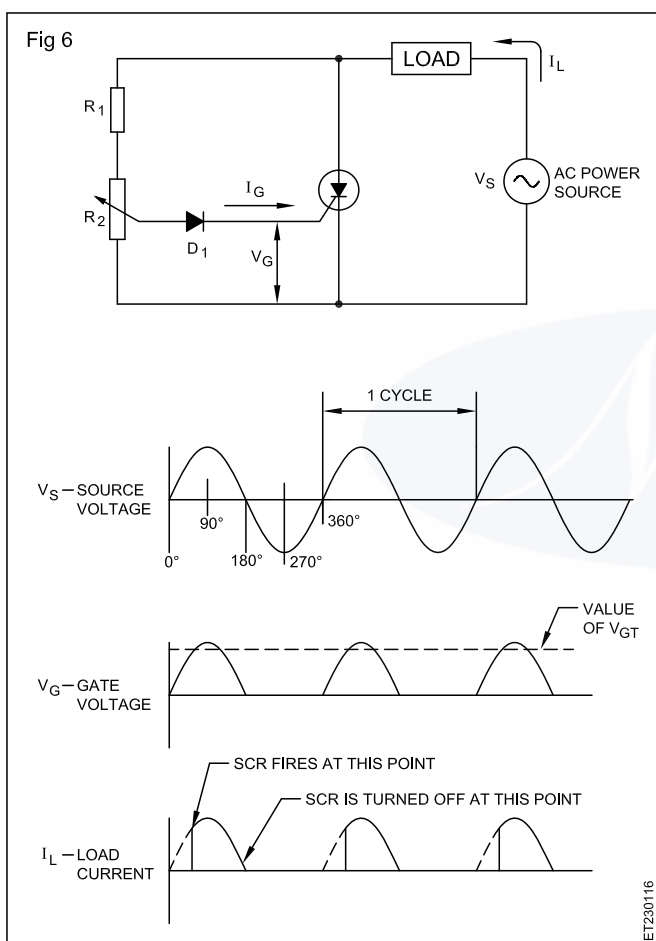


SCR operation in third quadrant with DC supply

When the anode-cathode of SCR is reverse biased a small reverse leakage current flows through the SCR. When the reverse bias voltage is increased slowly from zero to a higher value, beyond a particular voltage called the “Maximum Reverse blocking voltage”, the PN junctions of the SCR will break down and thus SCR becomes a short and enters into conduction. This results in a large current through the SCR/load as shown in fig 5 in the third quadrant. SCR should never be applied with a reverse voltage as high as this Maximum Reverse blocking voltage.

SCR operation with AC supply

Operation of SCR with AC circuit is similar to SCR operation in the first and the third quadrant as discussed in above paragraphs. Fig 6 illustrates working of SCR in AC control circuits.



The SCR gate circuit consists of resistor R1, potentiometer R2 and silicon diode D1. Resistors R1 and R2 act as a variable voltage divider. By adjusting the value of R2 the gate current I_G can be suitably modified. Diode D1 prevents negative voltage being applied to the gate when the ac supply is in the negative half cycle.

[X] During the +ve half cycle of the AC power source,

- as the positive half cycle voltage increases, the gate current I_G increases. When I_G reaches the trigger level, SCR fires and allows current I_L to flow through the load. From this point onwards the SCR impedance is low and

current I_L continues to flow throughout the +ve half cycle even though the gate current reduces below the trigger value (recall: once SCR is fired it continues to conduct even if the gate trigger is decreased or removed).

[Y] At the end of the +ve half cycle of AC power source,

- the +ve voltage drops to zero and SCR ceases to conduct (recall: one method of turning off SCR is to reduce the current through the SCR to below the holding current. This can be done by either opening the load circuit or reducing the supply to zero). Thus the SCR remains in off state throughout the negative half cycle.

Cycle [X] and [Y] repeats and current through the load flows in pulses as shown in Fig 6d.

Fig 6b, 6c shows the voltage wave forms of source and gate voltage.

If the value of R2 is varied, the point at which SCR triggers also varies changing the firing point shown in fig 6d. In the circuit shown in Fig 6a, the firing of SCR can be adjusted anywhere between almost 180 degrees (maximum) to 90 degrees (minimum).

This simple AC control circuit shown in Fig 6 using SCR can be used to control the current through the load during the +ve half cycle of AC. During the -ve half cycle the SCR remains turned off. Thus, SCR can be used as an excellent switching device in AC control circuits.

The circuit at Fig 6 is useful only in limited applications such as temperature control of soldering iron etc., More practical AC control circuits are discussed in further lessons.

SCR Specifications

For maximum performance, reliability and safety, each SCR should be operated within the ratings specified by the manufacturer. When you are designing a new circuit or replacing an SCR of an existing circuit, it is very essential to check the data of the SCR before using.

A typical SCR data sheet lists many specifications. A few of these specifications considered as most important from the point of view of a technician are listed below;

V_{DRM} : This is referred to as the anode-cathode forward break down voltage with no input gate current. Beyond this value, the SCR will break down into forward conduction. For typical SCRs, V_{DRM} ranges from about 30 Volts to 800 Volts.

V_{RRM} : This is referred to as the maximum anode-cathode reverse breakdown voltage. Beyond this value the SCR will break into reverse conduction. For typical SCRs, V_{RRM} also ranges from about 30 Volts to 800 Volts.

V_{GT} : This is the minimum positive gate voltage required to turn on the SCR. At temperature of 25°C typical values of V_{GT} is 0.7 to 0.8 volts. For higher temperatures of about 100 to 125°C, V_{GT} drops to approximately 0.2 volts.

I_{GT} : This is the minimum forward gate current required to turn on the SCR. Typically, low power SCRs require an I_{GT} of about 100 to 300 μA for turn-on. Medium and high power SCR needs about 5 to 150 mA for turn-on.

I_H : This is the minimum load current required to be maintained in the SCR to keep the SCR in on-state. Typical values of I_H ranges from about 6 mA for low power SCRs to 80 mA for high power SCRs.

I_T : This is the maximum allowable anode current that the SCR can withstand. This is defined by manufacturers in terms of rms forward current (I_{rms}) or average forward current (I_{avg}) for 180° of conduction. Typical values of I_{rms} for low and medium power SCRs range from about 1 to 30 Amperes.

V_{TM} : It is the maximum on state anode-cathode voltage drop across the SCR when it is conducting. Some manufacturers refer to this specification as V_F or V_{FM} . For most SCRs, V_{TM} is of the order of 1.6V.

t_{gt} : It is the time required for a specified gate current to turn on the SCR. Typical values lie in the range of 1 to 2 μ Sec.

t_q : It is the time required for a specified SCR to turn-off when the load circuit is made open. Typical values lie in the range of 15 to 35 μ Sec.

In addition to the above specifications, there are a few more specifications of SCR, some are of general nature to be considered while designing circuits using SCR and some are specific to certain SCR types. For further details on SCR specifications refer reference books listed at the end of this book or refer specification sheets of SCR supplied by the manufacturers.

Specifications of a few SCRs

2N5060

V_{DRM}	
V_{RRM}	30V
V_{TM}	
V_{GT}	1.2V
I_{GT}	0.35A t_{gt}
I_H	t_q
I_T	0.8A
Package type:	TO 92

MCR218-5

V_{DRM}	300V
V_{RRM}	
V_{TM}	1.5(typ) to 1.8V(max)
V_{GT}	0.2V(min) - 2.5V(max)
I_{GT}	10mA(typ) to 25mA(max)
I_H	16mA(typ) to 30mA(max)
$I_{T(rms)}$	8 Amps.
t_{gt}	
t_q	
Package type :	TO220

QUICK Check of SCRs

Quick check on SCRs can be carried out using a ohmmeter/multimeter. Since SCRs are made of PNPJ junction, resistance between junctions can be measured to conclude good working condition of the SCR. A good SCR shows following resistances between its terminal leads;

CHECK - 1

Between Anode - Cathode - Infinite resistance

[Irrespective of polarity]

Between Gate - Cathode

(i) Forward biased - Very low resistance
(30 to 500 ohms)

(ii) Reverse biased - High resistance

Between Gate - Anode - Infinite resistance

[Irrespective of polarity]

CHECK - 2

- Set multimeter to low resistance range.
- Connect positive lead of multimeter to the anode and the negative lead to the cathode. Meter should show infinite resistance.
- Now, for moment, short anode and gate of SCR by a piece of wire. The meter should show low resistance and will continue to show low resistance even after the short between anode and gate is been removed.

It is difficult to check leaky SCRs using only a ohmmeter/multimeter.

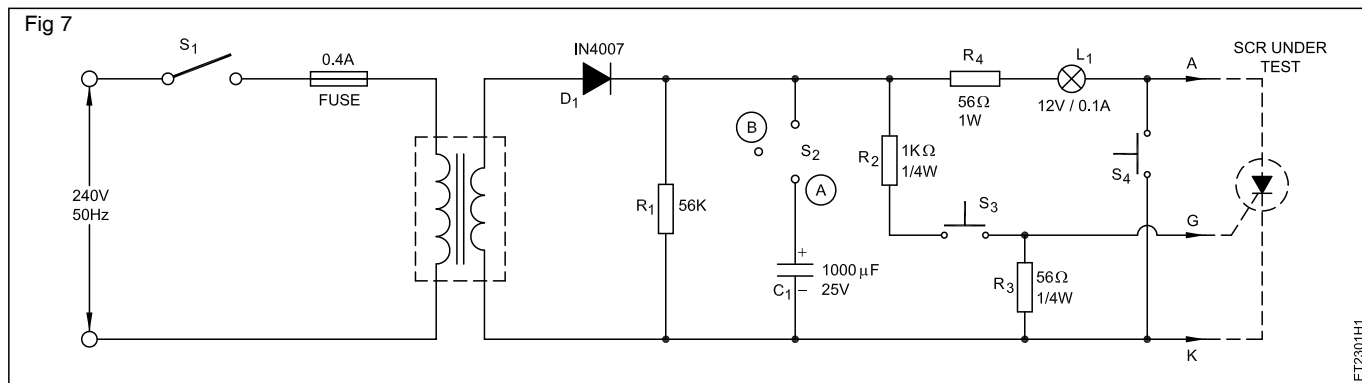
Checking SCRs using a SCR checking JIG

A simple "SCR checking jig" is shown in Fig 7. In Fig 7, the SCR shown in dotted lines is the SCR to be checked. This could be any SCR that you want to check.

In Fig 7, the stepped down to 12 AC is rectified by diode D_1 and filtered by capacitor C_1 when switch S_2 is at position A. The rectified and filtered DC is applied to the anode of SCR through lamp L_1 via a limiting resistor R_4 . Resistor R_2 , R_3 in series with a push button switch S_3 form a potential divider across dc. Voltage across R_3 is applied to the gate of SCR. This voltage is sufficient to turn on the SCR. Another push button Switch S_4 is connected directly across the Anode and Cathode of the SCR to be checked.

When a good SCR is placed in the test jig (in place of the SCR shown dotted) the circuit should function as given below;

1. When mains supply is on, and S_2 is at DC position, the lamp will start glowing as soon as the switch S_3 is pressed. Even on releasing S_3 , the lamp should keep on glowing. Under this condition, if switch S_4 is pressed once and released the SCR stops conducting and hence the lamp stops glowing.
2. If switch S_2 is put to AC position, lamp should glow only as long as switch S_3 is pressed.



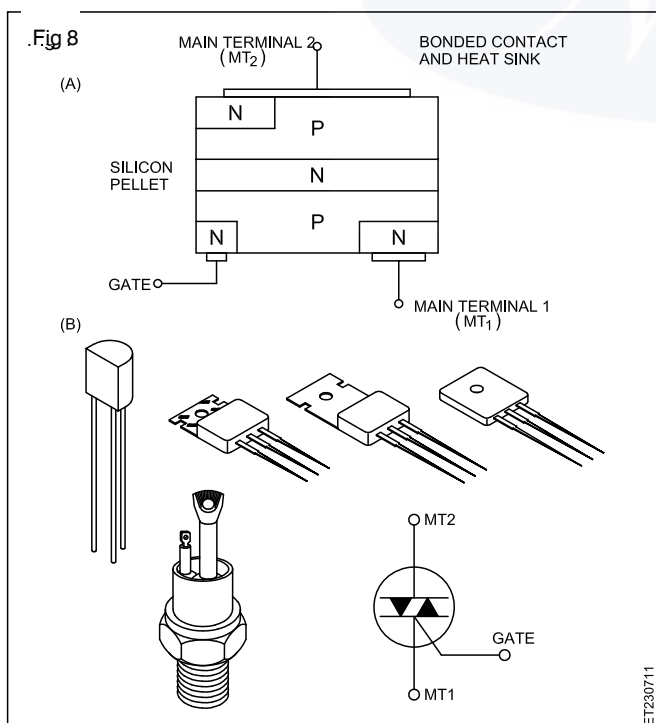
Instructor to discuss reasons for 1 & 2 above as an interactive discussion. Instructor should also discuss the outcome of above steps when a open or a shorted SCR is checked.

TRIAC

TRIAC is a three terminal gated semiconductor device for controlling a.c in either direction. The term TRIAC stands for TRIode AC semiconductor. TRIAC is very similar to that of two SCR connected in reverse parallel. A triac is able to conduct a large current in both directions, being triggered ON in one direction or the other by a gate pulse of the appropriate polarity.

Basic construction of a triac, its symbol and a typical Triac is shown in Fig 8a,8b and 8c.

As can be noticed in Fig 8, the electrodes of a triac are labeled as,



Main terminal-1(MT1)

Main terminal-2(MT2)

and Gate(G).

The terminals are so labeled because, this device operated in both directions and hence the terms anode and cathode does not apply.

TRIAC triggering

Triac can be triggered/turned-ON by,

- 1) applying a gate current,
- 2) exceeding the avalanche breakdown voltage V_{BO} ,
- 3) allowing the MT1 - MT2 applied voltage to increase at a rate in excess of the maximum dv/dt figure.

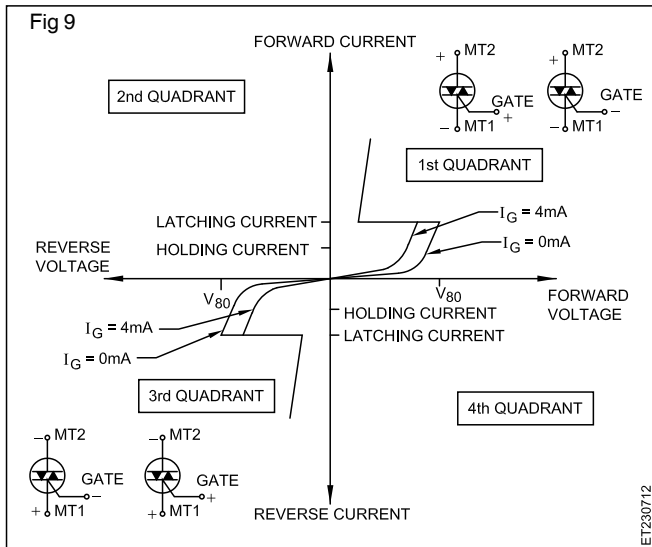
Methods 2 and 3 above are not employed in normal triac operation but they may be considered as limiting factors in circuit design. Hence all further discussion is restricted to triggering the triac via the gate. Since triac is a bi-directional device, it can be triggered into conduction by a negative or a positive gate signal. Triacs potentials are all considered with respect to main terminal-1(MT1). This gives the following possible *operating* situations or *modes*;

- MT2 +ve with respect to MT1 —Gate signal +ve (1st quadrant+)
- MT2 +ve with respect to MT1 —Gate signal -ve (1st quadrant-)
- MT2 -ve with respect to MT1 —Gate signal +ve (3rd quadrant+)
- MT2 -ve with respect to MT1 —Gate signal -ve (3rd quadrant-)

Unfortunately, triac is not equally sensitive in all the above said modes. It is least sensitive in 3rd quadrant mode(MT2 negative with respect of MT1 and trigger by a +ve gate signal) so this mode is very rarely used in practice.

When a triac is ON the current flowing between MT1 and MT2 is known as **Principal current**.

The triac will remain ON as long as the current flowing through it is larger than the holding current as shown in the static characteristics of a triac in Fig 9.



From the triac static characteristics. When MT2 is positive with respect to MT1, the triac operates in the first quadrant of its static characteristics, if it is not triggered, the small forward current increases slowly with increase in voltage until the breakdown voltage V_{BO} is reached and then the current increases rapidly. The device can be, and usually is, turned 'ON' at a smaller forward current by injecting a suitable gate current and the characteristics shows the effect of increasing the gate current from zero to 4 mA. The gate current must be maintained until the main current is at least equal to the latching current.

When terminal MT1 is positive with respect to MT2 the triac operates in the third quadrant and the current flows in the opposite direction.

FULL WAVE control using a TRIAC

Fig 10a shows a triac used for controlling the current flowing in an a.c circuit. Fig 10b shows the wave forms with different settings of POT V_{R1} .

On observing the waveforms at Fig 10b, it can be seen that control is achieved by firing the triac at the same point in both the positive and negative half cycle. Once triggered the device remains ON until the supply is switched OFF.

Choosing a TRIAC

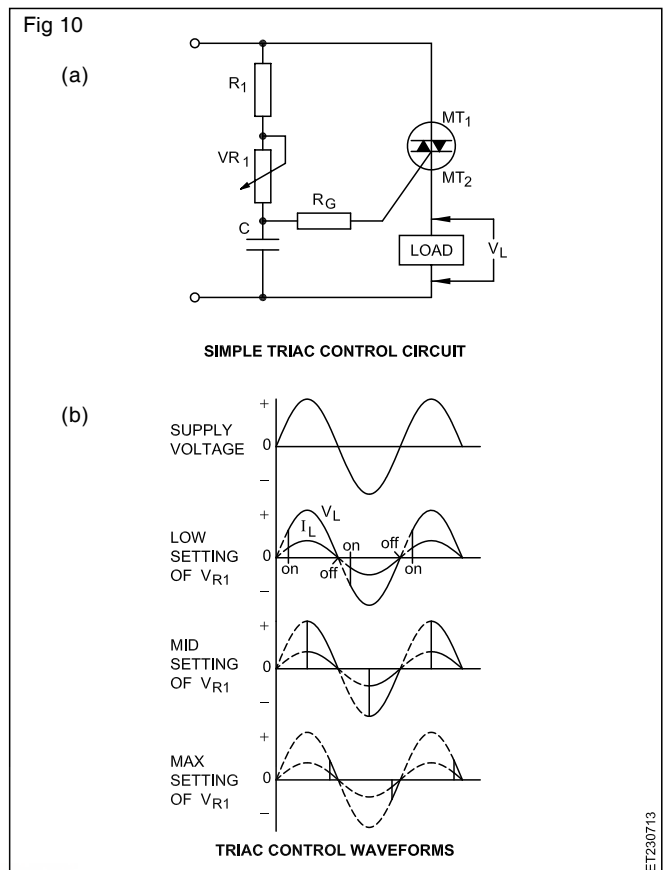
Like all other components, Triacs have maximum specified values of current and voltage that must not be exceeded. Important specifications of a triac with an example is given below;

TRIAC Type code : BT 136 TIC201D
 $I_T(\text{rms})$: 4 Amps. 2 Amps.

The maximum rms current that can continuously pass without damage occurring.

V_{GT} : 1.5 Volts. 2.5 Volts.

The value of gate current required to achieve switch on.



V_{DRM} : 400 Volts. 400 Volts.

The maximum permitted peak voltage.

NOTE: In triacs the terms forward and reverse do not arise since it is bidirectional.

QUICK TESTING TRAICS

A quick test can be carried out on Traic using an ohmmeter. If the readings taken are comparable to the one shown in table below, the Triac can be considered as satisfactory and can be used in circuit;

Meter polarities		Resistance
+	-	
MT2	MT1	>1M
MT1	MT2	>1M
MT2	G	>1M
G	MT2	>1M
MT1	G	$\approx 300W$
G	MT1	$\approx 300W$

Testing Triac DC Characteristics

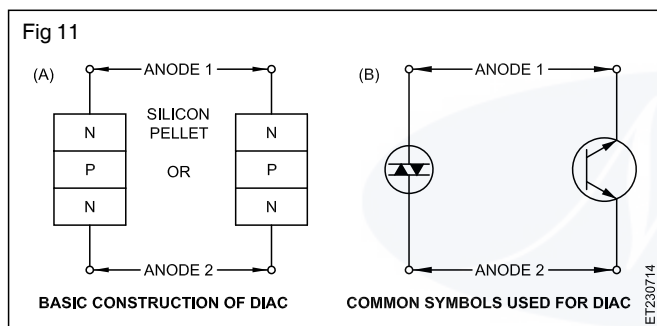
A simple test setup for basic triac characteristics is shown in Fig 1 of exercise 1.3.14. The purpose of this experiment is to investigate the switching characteristics and operation of a typical Triac with DC supply in different triggering modes.

Refer Fig 2 of exercise 1.3.14 while discussing the test circuit in class room.

The LED's used in the test circuit are to indicate the direction of current flow. With the gate voltage polarity in one direction, upon increasing the gate trigger voltage, triac fires and conducts in one direction and one of the LED glow. With the gate voltage polarity reversed, upon increasing the gate trigger voltage, triac fires and conducts in other direction and the other LED glows. In the experiment at 1.3.14 the values of V_{GT} and I_{GT} for a given Triac can be found in different modes of operation.

THE DIAC

Like UJTs, DIAC is a semiconductor device used extensively as a trigger device for thyristors gate circuit. In its most elementary form, DIAC is a three layer device as shown in Fig 11.

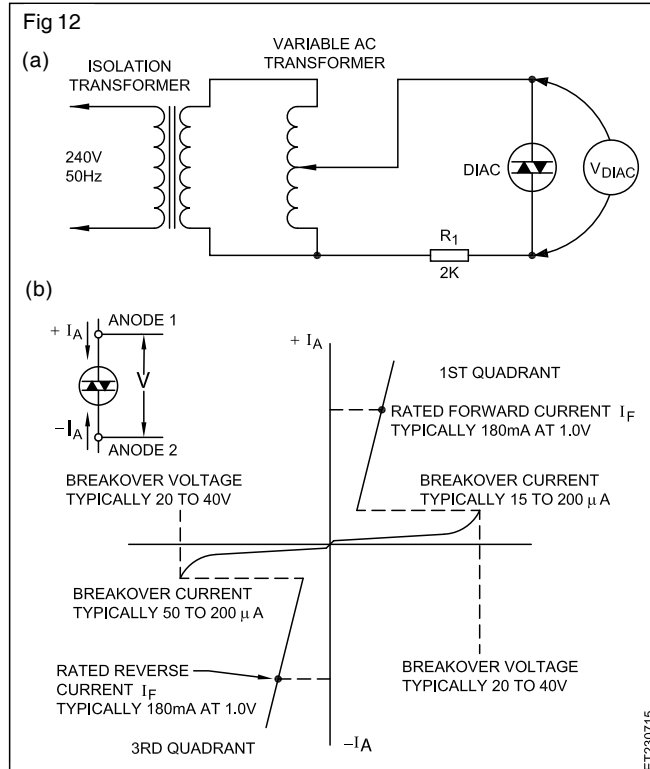


As can be seen from Fig 11, DIAC is a three layer, two terminal semiconductor device capable of conducting current in both directions.

DIACs resembles an NPN or PNP bipolar transistor with no base connection. Unlike bipolar transistor, the diac possesses uniform construction. This means, N-type and P-type doping is essentially the same at both junctions. As shown in Fig 11, diac may be constructed as either an NPN or PNP structure.

Fig 12a shows the experimental set up for testing the diac. The isolation transformer is used to isolate the circuit from the supply mains. The variable transformer or variation is used to apply the variable voltage to diac under test. The characteristic curve a typical diac is shown in Fig 12b.

As shown in the experimental set-up at Fig 12a, when a small voltage of either polarity is applied across a DIAC, the current flows is very small as can be seen from its characteristics in the first and the third quadrants. If the applied voltage is steadily increased, the current will



remain at a low value until the applied voltage reaches a value known as the Break over voltage of the DIAC as shown in Fig 12b. Once this point is reached the diac current increases rapidly and the diac voltage falls to a low value. At this point, the diac exhibits negative resistance characteristics (current conduction increases while the voltage across the device decreases). The diac will continue to conduct current as long as the current is greater than the Holding current of the device.

A diac acts in a similar manner to two diodes that are connected in reverse parallel and it therefore is able to rectify a.c voltage during both half cycles. The symbol used for DIAC is shown in Fig 11b.

Choosing DIACs

Quick Testing DIACs

Since Diacs are similar to two diodes connected back to back and break down in either direction once the applied voltage reaches the breakdown voltage of the diode, while testing a diac using a ohmmeter, it should show high resistance (infinite resistance) when checked in either direction. This quick test only confirms that the DIAC is not shorted; however this quick test is worth carrying out before using the Diac in a circuit.

Crowbar protection using SCR(s)

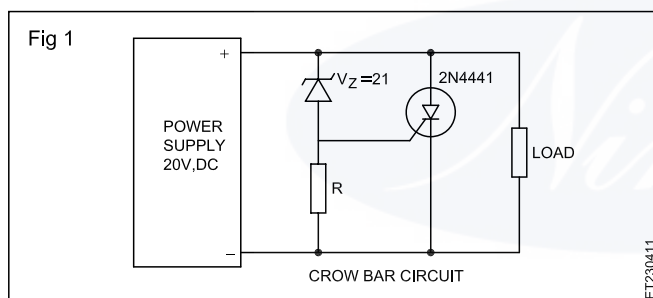
Objectives: At the end of this lesson you shall be able to

- explain with a sketch crow bar protection circuit using SCRs in DC Power supplies.
- explain with a sketch over voltage protection using SCRs for AC mains supplies.
- explain with a sketch static circuit breaker using SCRs
- explain the use of SCR in ac and dc power control
- explain the meaning and use of time delay circuits
- explain with an example the need and use of soft turn-on circuits
- explain which property of SCR makes it usable in logic and digital circuits.

Over voltage/Crow bar protection in power supplies using SCR

One of the application of SCRs is to protect a load against over voltage. As an example, in circuits using TTL digital ICs, if the dc supply voltage to these circuits exceeds 5 volt the digital ICs get damaged. To prevent such damages, a crow bar circuit is used. Essentially a crowbar circuit short the load terminals at the first sign of over voltage.

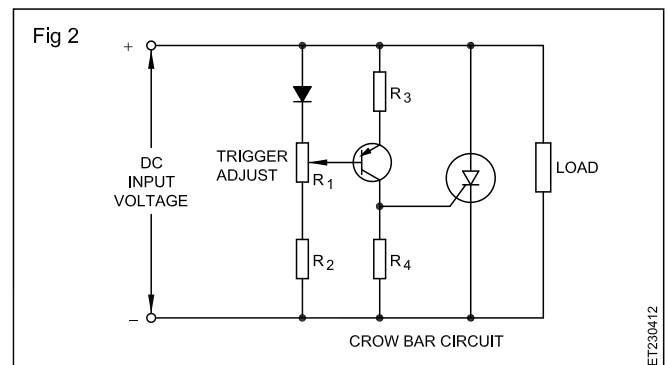
Fig 1 shows the principle of crow bar protection. In this a 20V dc supply is connected to the load through a crow bar protection.



The 20V dc positive supply is connected across a zener diode in series with a resistor. The voltage across the resistor drives the gate of an SCR which is in parallel with the load. The zener voltage is 21 volts. As long as the supply voltage is 20 V the zener diode is open and there is no gate current and hence the SCR is open. Since the SCR is open, 20 volts appears across the load.

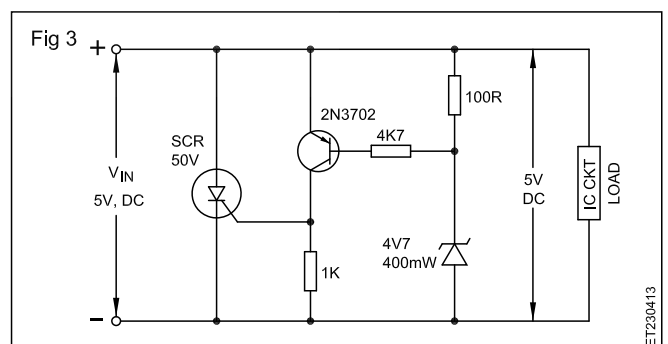
If the DC supply voltage happens to rise to 21 volt and above, the zener diode breakdown, current flows through the zener, a voltage drop appears across the resistor which in turn appears at the gate of the SCR. This gate voltage makes the SCR to conduct making the load voltage almost equal to zero. This action is similar to throwing a 'crowbar' across the load terminals. Because the SCR turn-on is very fast, if the input supply voltage increases to a level which may damage the load, the load is quickly protected by reducing the load voltage to zero.

Fig 2 shows a practical SCR crowbar incorporating a transistor which provides sufficient voltage gain for a more sharp turn-on compared to the one shown in Fig 1.

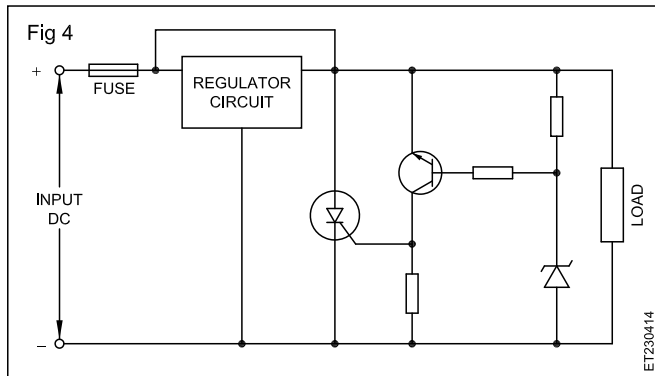


When the voltage across R_4 exceeds approximately 0.7 volts, the SCR turns on. An ordinary junction diode is included for temperature compensation of the transistor base-emitter diode. The trigger adjust allows to set the trip point of the circuit typically around 10 to 15 percent above the expected output voltage.

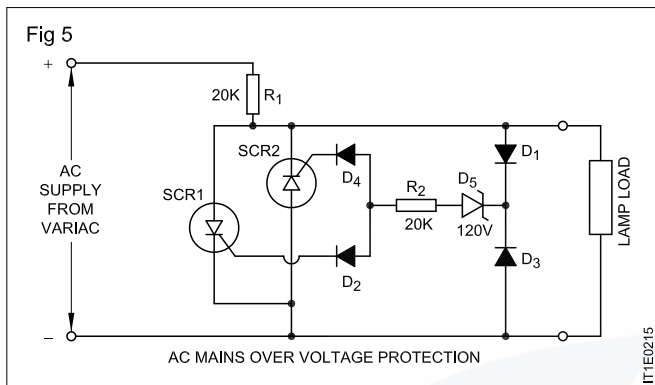
Fig 3 shows a practical Crowbar circuit used for 5V supply voltage.



In the case of a supply with only fuse protection it is advised to connect the thyristor across the unregulated supply as shown in Fig 4. This arrangement prevents the possible damage to the regulator circuit when the crow bar operates.



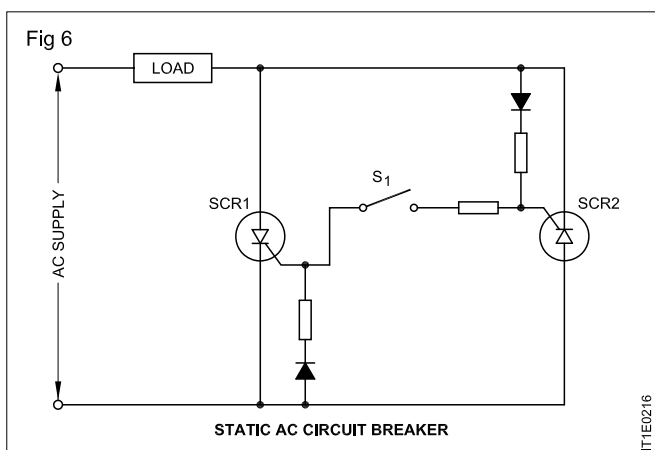
A over voltage protection circuit used with AC mains supply is shown in Fig 5. Since the applied voltage is AC, as shown in Fig 5, two SCRs are used one for the +ve half cycle and the other for the negative half cycle.



When the AC mains supply voltage is rises above the specified limit, then, during the positive half cycle, the gate of SCR1 will get energized through R1,D1,R2,D2 and trigger SCR1. During the negative half cycle, the gate of SCR2 will get energized through D3,R2,D4,R1 and SCR2 will turn on if the over voltage persist. As soon as the load voltage returns to the safe value, the zener diode D5 will recover and the current through it will be very low so that the SCRs will not fire. The response time of the protection circuit is fast and the protection will be provided in both positive and negative half cycles of the ac input. In the circuit at Fig 5, resistor R1 is used to limit the short-circuit current when the SCRs are fired.

Static circuit breaker using SCRs

Fig 6 shows a simple AC circuit breaker circuit. In this circuit SCRs are used for making and breaking the circuit.



When the switch S1 is closed, SCR1 will fire at the beginning of the +ve half cycle. SCR1 will turn-off when the positive half cycle goes through value. As soon as SCR1 is turned off, SCR2 will fire since the input ac voltage polarity reversed.

To break the circuit, switch S1 is opened. Although the current through the load circuit may be heavy, the current through the gate circuit and switch is very small, therefore opening the gate circuit poses no problem. With the switch S open, no further gate signal will be applied to the SCRs and thus the SCRs will not be triggered and hence the load current will be zero.

The maximum time delay for breaking the circuit is one-half cycle.

The circuit shown in Fig 6 is one type of circuit breaker. There are several possible variations to it. Refer reference books listed at the end of this book for further reading.

Power control circuits

The bistable property(ON/OFF) of SCRs and the efficient gate control to trigger the SCR makes it ideally suited for many industrial applications. One such application is in the control of AC and DC power delivered to the load. Such power control is used in voltage control of DC power supplies, temperature regulators, light dimmers etc., Details of power control application circuits are discussed in further lessons as majority of these circuits use a few more components not so far discussed in previous lessons.

TIME DELAY CIRCUITS

Whenever power is to be applied or removed from a circuit at a pre determined time after initial signal applied in industrial circuits, time delay circuits are frequently used. SCRs in conjunction with a few other components can be used to provide the necessary time delays. Time delay circuits are discussed in detail in further lessons after introducing a few other components that go along with the circuit.

SOFT TURN-ON CIRCUITS

In many power control circuits, it is desirable to apply power gradually rather than suddenly. For example in the lamps used with overhead projectors, the cold resistance of the lamp filament is low and there will be a very high current surge if full voltage is applied at once. This heavy surge current may damage the costly lamps used in overhead projectors. By using a soft start circuit, using SCR and using phase control, voltage is applied gradually to the lamp protecting the lamp from getting damaged.

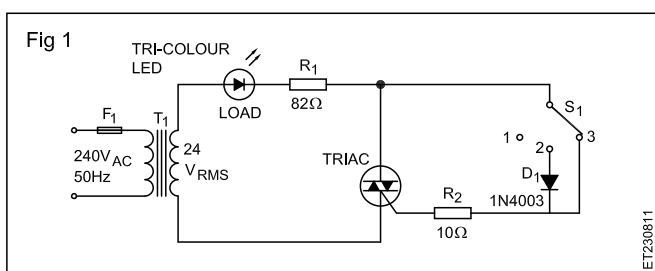
AC power control using TRIAC

Objectives: At the end of this lesson you shall be able to

- explain the working of Triac in half wave and fullwave a.c operation
- explain working of Triac with suitable gate trigger circuit
- explain the use of Triac as a typical security alarm system
- explain working of a simple unattended porch/night lamp using LDR and Traic.

TRIAC as a three-way a.c power switcher

The circuit at Fig 1 using a Triac can be used to provide a three way power switching capability.



In Fig 1 as above,

- With switch S1 in position-1, power to the load is OFF.
- With switch S1 in position-2, power to the load is HALF(Half-Wave).
- With switch S1 in position-3, power to the load is FULL(Full-Wave).

This allows power to the load to be adjusted to, No-load, Half-load and Full-load conduction. Also, by increasing or decreasing the value of R_2 , the firing angle can be delayed to provide one-fourth load to approximately full-load conditions.

Further

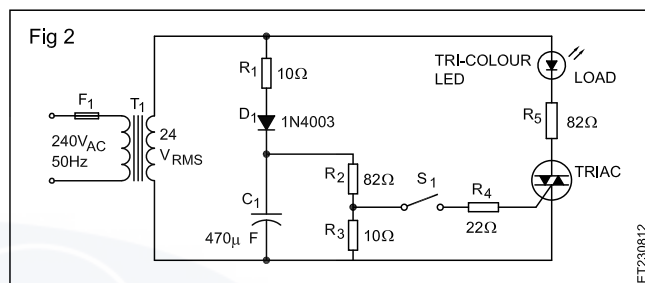
- With switch S1 in position-1, there is no gate drive, hence the triac is off and therefore the load current is off.
- With switch S1 in position-2, triac conducts only during positive half cycles and therefore the load current is only in one direction of a.c. The tri-colour LED glows red indicating flow of current in only one direction through the time.
- With switch S1 in position-3, triac conducts during both +ve and -ve half cycles and therefore the load current is full. The tri-colour LED glows yellow indicating flow of current in both directions through the triac.

Refer lesson 1.3.16 for details on multicolour LED's.

If the value of R_2 is increased such that there is a delayed gate triggering, the triac can be made to conduct only during half the time of the positive or negative half cycle such that with the condition listed at [b] above, there is only one-fourth of the load current instead of one-half.

The circuit at Fig 1 although is simple and reliable, this circuit introduces radio-frequency-interference(RFI) resulting in AM disturbance. This RFI is a serious disadvantage of this circuit. The RF interference can be verified by placing a AM radio receiver near the triac switching circuit with the switching circuit operational.

The disadvantage a.c switching circuit at Fig 1 is overcome in the control circuit shown in Fig 2.

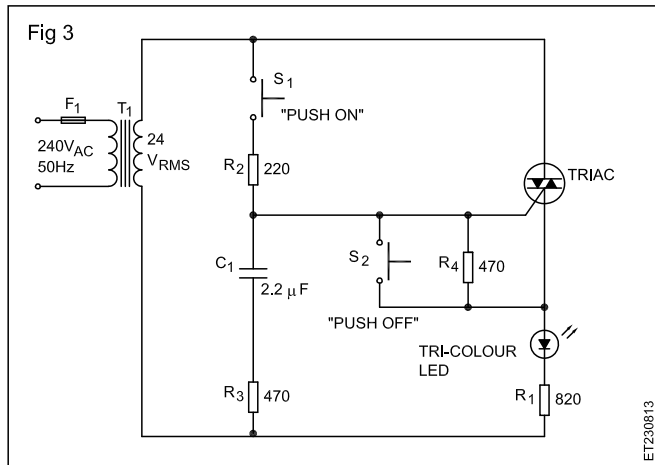


The RFI generated in the triac power control circuit shown in fig 1 can be virtually eliminated by using gate trigger circuits that turn-on the triac at the beginning of the a.c power cycle. For example, a constant d.c gate trigger voltage allows the triac to be switched-on at the beginning of the a.c cycle; this eliminates any abrupt switching which generates rfi radiation. Fig 2 shows how this d.c trigger voltage may be developed from an a.c power source. This circuit can be tested as explained earlier to confirm for no RFI. Note that even this circuit provides power control for No-load, Full-load or intermediate load conditions.

Security alarm systems using Traic

In certain situations, such as security alarm systems, momentary contact is required to turn on or off a triac control circuit. Such security alarm systems use the principle of self latching to detect unauthorized intrusion. Fig 3 illustrates a self latching type security alarm system concept.

When switch S1 is momentarily closed, the initial gate-switching pulse is provided through R2 and the triac is turned-on. As the triac is turned on, C1 is charged during one of the a.c half-cycles. C1 starts to discharge through R3 when the triac is turned-on. However, the slow time constant of R3/C1 maintains a voltage to the triac gate terminal after the a.c power half-cycle, this turns on the triac during the next a.c half-cycle. At the same time, the capacitor C1 is recharged for the next a.c half-cycle. Thus, the triac continues to be turned on as it is self-latching.



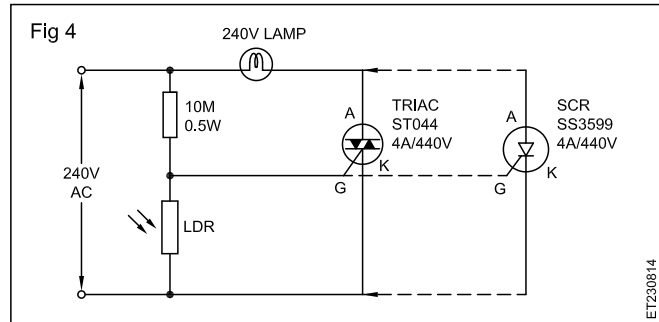
In order to turn-off the triac, Switch S2 is closed momentarily. Closing S2 has the effect of shorting the triac gate terminal to load terminal MT₁. Thus, a zero gate voltage condition causes the triac to be turned off.

The circuit at Fig 3 can be used as a security alarm system. Security alarm systems use either SCRs with d.c power supplies, or Triacs with a.c power supplies. The d.c powered security alarm systems usually have a battery for sustained operation in the event the mains a.c power failure.

Simple unattended porch light/Night lamp using Traic

A simple circuit using a Light Dependent Resistor(LDR) and a Triac which can be used for applications such as stair case light, Gate light etc., where unattended illumination is required after sunset is shown in Fig 4.

In the circuit shown in Fig 4, as the day light fades, the LDR resistance increases and hence the voltage across the LDR becomes sufficient enough to trigger the triac in to condition. After sun rise the resistance of LDR decreases



and the voltage across it becomes insufficient to trigger the triac into conduction and hence the lamp turns-off.

The precaution to be taken while installing this circuit is that, once the lamp turns-on, the light from the lamp that is turned on should not fall on the LDR.

If the lamp used for light is less than 20 watts, there is no need of a heat sink for the triac. For lamp power ratings above 20 watts it is necessary to use a adequate heat sink for the triac.

In the circuit at fig 4, SCR can also be used in place of a triac; but the lamp will give only around 70% of light. This reduced light can be compensated by using a higher wattage lamp.

Lamp dimmer/fan motor speed regulator using TRIAC and DIAC

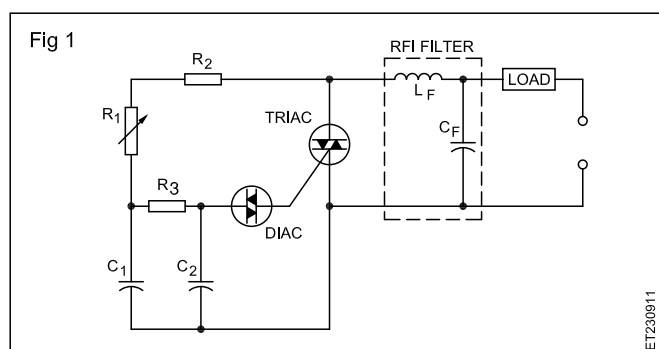
Objectives: At the end of this lesson you shall be able to

- know the advantage of use of TRIAC for speed control of a.c motors
- know the method of making motor speed control repeatable
- explain the need of soft-start light dimmers
- draw circuit of a simple light dimmer using a diac and a triac
- explain the function of components of a typical soft-start light dimmer
- explain the need and working of snubber circuit in a light dimmer circuit.

TRIAC or SCR for speed control of a.c motors ?

Compared to SCR, Triac is more suitable for speed control of universal motors. Although both SCR and TRIAC can be used to phase control and vary the current through the motor, Triac being a full-wave device, symmetrically controls the phase of both half cycles of the applied ac. The resultant full-wave current format then produces smoother motor operation than can be attained from the half-wave rectification using SCRs. This is particularly noticeable during speed control at low speeds.

The circuit at Fig 1 shows a Triac phase-control circuit for controlling the speed of universal motors.



The load shown in circuit at Fig 1 is a general load rather than a motor symbol because, this circuit can also be used for light dimmers and for the control of heaters.

This circuit features a double time constant phase-shift network. This reduces hysteresis in firing of the triac, thereby making the manual adjustment of speed more repeatable.

The diac used as trigger device, adds to the reliability of the circuit. The elemental low-pass filter comprising L_F and C_F attenuates much of the radio-frequency interference(RFI)) that gets generated and tries to get in to the power line. Such high-frequency RFI energy is generated by the extremely rapid turn-on time of the Triac.

LAMP DIMMERS

Lamp dimmer is a circuit which controls ac power supplied to an incandescent lamp thereby controlling the intensity of light emitted by the lamp from almost zero to full brilliance.

Conventional and soft-start dimming of incandescent lights

Old technology light dimmers used high wattage rheostats, adjustable auto-transformers, or saturable reactors, which were large, expensive, and generated considerable heat. Present day semiconductor light dimmers have over come these deficiencies and have therefore become very popular for many applications.

Modern thyristor dimmers are inexpensive, reliable, small, generate little heat, and are easy to control remotely. These properties have not only permitted semiconductor dimmers to supersede older types in theaters and auditoriums with excellent results, but have made dimmers practical for built-in home lighting, table and floor lamps, projection equipment and other uses.

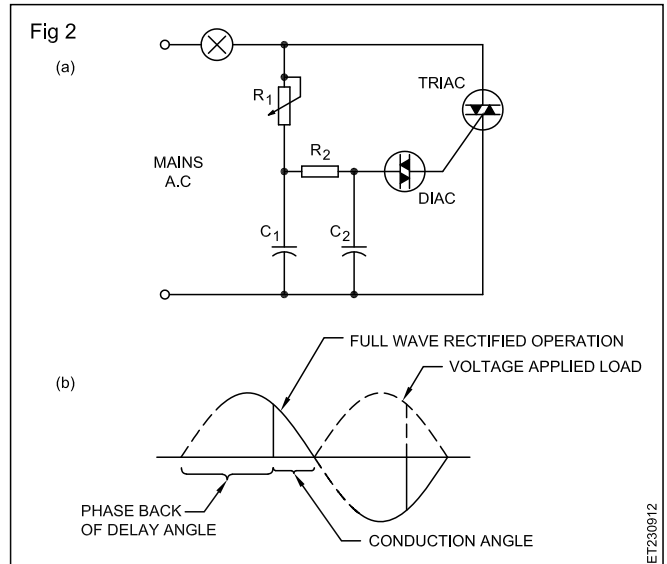
Two light dimmers for incandescent light bulbs are discussed below. Both these dimmer circuits control light intensity by adjusting the angle of conduction of a triac connected in series with the bulb. The first dimmer uses a very simple circuit that is ideal for highly compact applications requiring minimum cost. The second dimmer features soft starting for low inrush current and consequent long lamp life. Soft start lamp dimmers are especially useful with expensive lights with short lives, such as projection lamps and photographic bulbs.

Simple light dimmer

The circuit shown in Fig 2 is a wide range light dimmer using very few parts. The circuit can be operated using any mains supply source(240V, 50 Hz) by choosing appropriate value of circuit components. The circuit can control up to 1000 watts of power to incandescent bulbs.

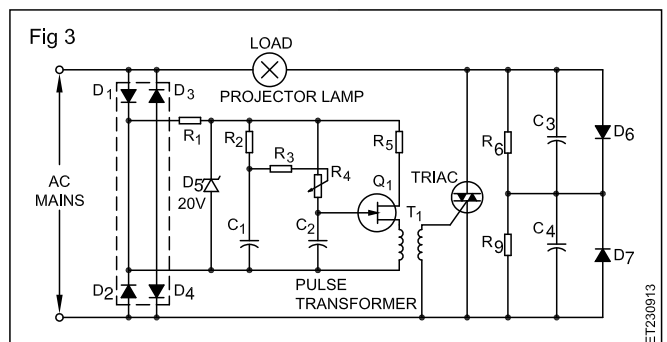
The power to the bulbs is varied by controlling the conduction angle of triac. Many circuits can be used for phase control, but the single triac circuit used is the simplest and is therefore chosen for this particular application.

The control circuit for this triac must function as shown Fig 2b. The control circuit must create a delay between the time voltage is applied to the circuit and the time it is applied to the load. The triac is triggered after this delay and conducts current through the load for the remaining part of each alternation. This circuit can control the conduction angle from 0° to about 170° and provides better than 97% of full-power control.



Light Dimmer with soft-start option

The circuit at Fig 3 is a light dimmer with soft start option. Soft starting is desirable because of the very low resistance of a cold lamp filament compared to its hot resistance. At the time of initial switching on, the low resistance of the lamp causes very high inrush currents which leads to short filament/lamp life. Lamp failures caused by high inrush currents is eliminated by the soft start feature, which applies current to the bulb slowly enough to eliminate high surges.



Operation of the circuit at Fig 3 begins when voltage is applied to the diode bridge consisting of D1 through D4. The bridge rectifies the input and applies a dc voltage to resistor R1 and zener diode D1. The zener provides a constant voltage of 20 volts to unijunction transistor Q1, except at the end of each alternation when the line voltage drops to zero. Initially the voltage across capacitor C1 is zero and capacitor C2 cannot charge to trigger Q1. C1 will begin to charge, but because the voltage is low, C2 will have adequate voltage to trigger C1 only near the end of the

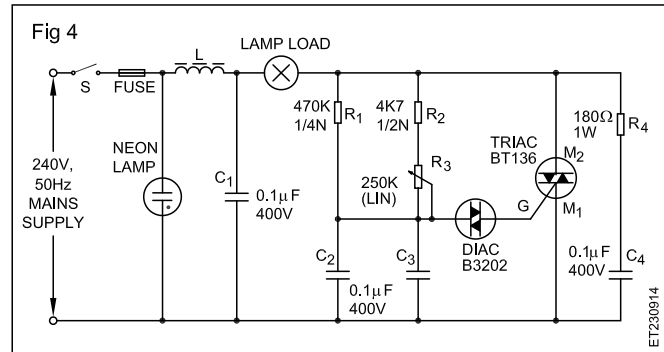
half cycle. Although the lamp resistance is low at this time, the voltage applied to the lamp is low and the inrush current is small. Then the voltage on C1 rises, allowing C2 to trigger Q1 earlier in the cycle. At the same time the lamp is being heated by the slowly increasing applied voltage and by the time the peak voltage applied to the lamp has reached its maximum value, the bulb has been heated sufficiently so that the peak inrush current is kept to a reasonable value. Resistor R4 controls the charging rate of C2 and provides the means to dim the lamp. Power to the load can be adjusted manually by varying the resistance of R4. T1 is a pulse transformer. In addition to supplying the trigger to Triac, this transformer isolates the high current load circuit from the low power triggering circuit (gate isolation methods for Triac in discussed in further paragraphs).

A simple Lamp dimmer cum Universal motor speed controller

In the lamp dimmer cum universal speed controller circuit shown in Fig 4, a triac is used as control device. Phase control technique is used to control conduction angle of the triac which in turn control the power fed to the lamp.

A lamp L is connected in series with a.c mains supply to the triac. The trigger pulses to triac gate is given through diac. The diac is triggered at the same break over voltage level (30 V) during both positive and negative half cycles.

Potentiometer R₄ provides the facility for varying the intensity of light or speed of a universal motor.



SNUBBER CIRCUIT

One problem with the triac control is the sudden application of reverse voltage across the triac immediately after it has stopped conduction. This is a serious problem when the load is a highly inductive as in motors. This reapplied voltage denoted by dv/dt can trigger-on (unwanted or false triggering) the device losing the phase control.

To avoid this false triggering, an R and C series network is placed across the circuit (R₅ and C₄ as shown in Fig 4). This RC network slows down the rate of rise of voltage applied across the triac. This RC circuit connected across the triac circuit is called Snubber circuit.

The inductance L and capacitor C₁ forms a low pass filter to substantially reduce the radio frequency interference (RFI) generated by the rapid turn-on and turn-off the triac.

The lamp dimmer circuit at Fig 4 can be used equally well as a fan speed regulator. The only change to be made is to connect a fan in place of the lamp shown in the circuit at Fig 4. The speed can be varied from almost zero to full speed by just rotating POT R₃.

Series voltage regulators

Objectives: At the end of this lesson you shall be able to

- state the disadvantages of a zener regulator
- state the working of a simple series regulator
- design a simple series regulator for a required dc output voltage
- write the blocks of a series regulator with voltage feedback and explain their functions
- design a series regulator for a given input-output specification
- state the modification necessary to make the output of a voltage regulator variable
- state the characteristics of a darlington pair connection
- state the method of increasing the load current capacity of series regulator.

Disadvantages of a zener regulator

Voltage regulated power supply using zener diode discussed in Lesson 1.3.17 is the simplest form of voltage regulator. But, zener voltage regulators have two main disadvantages:

- 1 When the load current requirement is higher, say of the order of a few amperes, the zener regulator requires a very high wattage zener diode capable of handling high current.
- 2 In a zener regulator, the load resistor sees an output impedance of approximately the zener impedance, R_Z which ranges from a few ohms to a few tens of ohms (typically 5Ω to 25Ω). This is a considerably high output impedance because the output impedance of an ideal power supply should be zero ohms.

These two disadvantages of zener regulators are overcome in a simple series regulator shown in Fig 1a.

The simple series regulator shown in Fig 1a, redrawn in Fig 1b is nothing but a zener regulator followed by an emitter follower. A circuit like this can hold the load voltage almost constant, thus working as a voltage regulator.

The advantages of this circuit are listed below;

1 Less load on the zener diode.

In this circuit Fig 1a or 1b the base current is very much smaller than the emitter current or the load current, a very small wattage zener diode itself is sufficient.

For instance for a load of say 1 amp, if the β_{dc} of the transistor is 100, then the zener diode need to handle only, 20mA for a typical $I_{Z(min)}$ of 5 to 10mA.

2 Lower output impedance

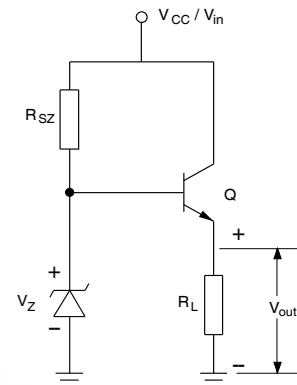
This circuit also offers a lower output impedance of $0.07W$ which is close to the ideal output impedance of zero required for a power supply. For further information, refer IMP of Electronic Mechanic 1st year.

Working of a simple series regulator

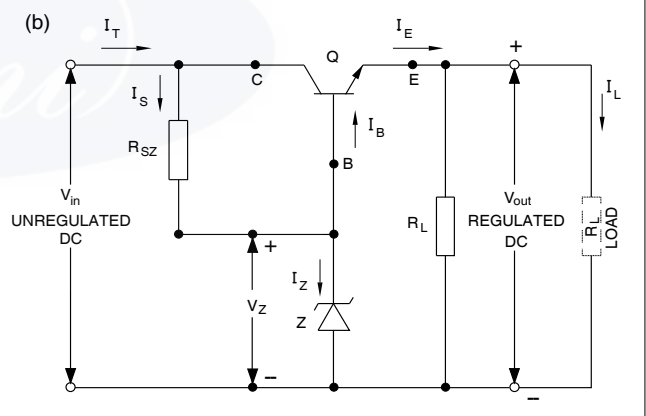
In Fig 1b, the current through R_{SZ} should be at least equal to zener breakdown current, plus, base current for the transistor Q.

Fig 1

(a)



(b)



The voltage across the zener, V_Z drives the base of the emitter follower. Therefore, the dc output voltage is bootstrapped to within one V_{BE} drop of the zener voltage. The regulated dc output voltage will be,

$$V_{out} = V_Z - V_{BE} \quad \dots\dots[1]$$

The collector-emitter voltage across the transistor will be the difference in the voltage between the input and output.

$$V_{CE} = V_{in} - V_{out}$$

If the input voltage V_{in} increases, the output voltage V_{out} remains constant due to the bootstrapped zener voltage. Therefore, the drop across the collector-emitter, V_{CE} of the transistor increases compensating the rise in the input voltage V_{in} .

For example, in the series regulator shown in Fig 1b, if V_{in} is 15 V and V_{out} is 12 volts, then, V_{CE} will be,

$$V_{CE} = V_{in} - V_{out} = 15 - 12 = 3 \text{ V.}$$

If V_{in} increases to say 20 V, then V_{CE} increases to $20 - 12 = 8$ V, thus keeping the output voltage unaltered at 12 volts. Since the collector and emitter of the transistor in Fig 1b is in series with the input and output terminals, and carries all the load current and this type of regulators are known as *series voltage regulators*. Hence the transistor is referred to as the *pass transistor*.

Because of the fact, that all the load current must flow through the pass transistor and that the value of V_{CE} increases when V_{in} increases, the wattage rating of the pass transistor should be high enough to handle the dissipation.

For instance, while supplying a load current of 300 mA, with V_{in} at 20 V and V_{out} at 12V, V_{CE} will be 8 V. Therefore, the dissipation at the transistor will be,

$$P_D = V_{CE} \times I_L = 8 \times 300 \text{ mA} = 2400 \text{ mw} = 2.4 \text{ watts}$$

To accommodate this, the wattage rating of the chosen pass transistor should be greater than 2.4 watts.

TIP: Allow at least 20% higher rating. For the example above, choose a transistor of rating $2.4 + 0.48$ watts 3 watts.

Because there will be quite a high dissipation depending on load current requirement, medium to high power transistors are used as pass transistors.

Temperature effect on output voltage

When temperature increases, V_{BE} decreases. Therefore, V_{out} increases by the change of current in V_{BE} .

Data sheets of transistors usually give information about how much V_{BE} changes with temperature.

For all practical purposes, an approximate of 2 mV decreases in V_{BE} for each degree rise in temperature. For instance when the temperature of the transistor rises from 25°C (room temperature) to 75°C (due to power dissipation at the transistor), V_{BE} decreases approximately 100 mV. Hence, the output will increase by 100 mV. This is relatively small, and hence, can be neglected.

Temperature also has an effect on the voltage across the zener. Any increase or decrease in the voltage across the zener is reflected at the output. Hence, while choosing the zener, it is equally important to know the temperature coefficient of the zener, specially when the power supply is connected to higher loads of the order of a few amperes.

Design guidelines for a simple series regulator

Example: For the series regulator shown in Fig 1, find the values of the components for an output voltage of 12 volts. The input unregulated dc voltage V_{in} to the regulation may vary 15V to 18V. The load resistor R_L is 220Ω.

Step 1

Finding zener voltage rating

Required output voltage $V_{out} = 12$.

$$V_{out} = V_Z - V_{BE}$$

$$\text{or, } V_Z = V_{out} + V_{BE} = 12 + 0.7 = 12.7 \text{ V}$$

Since the nearest value is 12 V, choose a zener of voltage 12 V.

A zener of 12 V, $\pm 5\%$ rating will have output of $12 \text{ V} + 0.6 \text{ V} = 12.6 \text{ V}$. Hence practically the output voltage of the regulator will be $12.6 - 0.7 \text{ V} = 11.9$. Hence for further calculations V_{out} is taken as approximately 12 V.

Step 2

Finding the value of R_{SZ} and the wattage of the zener diode.

a) At the given, the worst case (minimum) V_{in} of 15V, the minimum current through the zener to be in breakdown condition $\approx 10 \text{ mA}$.

$$R_{SZ} = \frac{15 - 12}{10 \text{ mA}} = \frac{3}{10 \text{ mA}} = 300 \Omega \approx 330 \Omega$$

b) When $V_{in} = 18 \text{ V}$ (given)

$$I_{SZ} = \frac{18 - 12}{330 \Omega} = 18.2 \text{ mA}$$

Therefore the wattage of zener should be,

$$P_Z = V_Z \times I_{SZ} = 12 \times 18.2 \text{ mA} = 218 \text{ mW.}$$

Hence, a 1/4 watts, 12V zener is sufficient.

Step 3

Finding wattage rating of pass transistor.

a) For a load of 220W, the load current will be,

$$I_L = \frac{V_{out}}{R_L} = \frac{12}{220 \Omega} = 54.5 \text{ mA} \approx 55 \text{ mA.}$$

b) Maximum power is dissipated in the transistor, when

(i) V_{in} is maximum, i.e. $V_{in} = 18 \text{ V}$.

and (ii) Load current is maximum, i.e. $I_L = 55 \text{ mA}$.

Therefore, Power rating of pass transistor P_Q should be,

$$\begin{aligned} P_Q &= V_{CE} \times I_L \\ &= (V_{in} - V_{out}) \times I_L \\ &= (18 - 12) \times 55 \text{ mA} \\ &= 6 \text{ V} \times 55 \text{ mA} \\ &= 330 \text{ mwatts} \end{aligned}$$

Choose any transistor with a power rating of 20% more than P_Q i.e. $330 + 66 = 396 \text{ mW} \approx 400 \text{ mW}$

For further calculations SL100 is chosen as the pass transistor although its wattage rating is less than 1 watt. The damage to the transistor can be avoided if load current is restricted to <28mA.

c) The base current for the chosen pass transistor

SL100/HL100 has β_{dc} between 50 to 200. Assuming an average β_{dc} of 80,

$$I_B = \frac{I_E}{\beta_{dc}} = \frac{55 \text{ mA}}{80} = 0.69 \text{ mA}$$

Total maximum zener current I_Z will be, $I_Z = I_{SZ} - I_B$

$$I_Z = 18.2 \text{ mA} - 0.69 \text{ mA} \approx 17.5 \text{ mA}$$

Step 4

Finding the output impedance of the regulator.
Assuming the zener resistance, $R_Z = 7\Omega$

$$Z_{out} = r'_e + \frac{R_Z}{\beta_{dc}}$$

where,

$$r'_e = \frac{25 \text{ mV}}{I_E} = \frac{25 \text{ mV}}{55 \text{ mA}} = 0.46\Omega$$

and, $\frac{R_Z}{\beta_{dc}} = \frac{7\Omega}{80} = 0.0875$

Therefore,

$$Z_{out} = r'_e + \frac{R_Z}{\beta_{dc}} = 0.46 + 0.0875 = 0.548\Omega$$

r'_e is high because I_L is low. If I_L is high then r'_e becomes very small.

Step 5

Finding V_{out} variations for higher loads. If the load resistance is decreased from 220Ω to 180Ω , the load current increases from 55 mA to 67 mA. This change in 12 mA of load, decreases the load voltage by,

$$DV_L = (DI_L)(Z_{out})$$

Therefore,

$$DV_L = (12 \text{ mA})(0.548 \text{ W}) = 6.58 \text{ mV}$$

This change is negligible compared to 12V and hence, for all practical purposes, the output voltage can be taken as well regulated.

Power supply characteristics

A power supply is said to be good, if its output voltage changes by as small a value as possible when the load current changes or when the input voltage changes. These are specified in terms referred as,

- Load regulation
- Source regulation [also called line regulation.]

Load regulation

Load regulation is defined as the change in the regulated output voltage when the load current changes from the minimum to the maximum.

$$\text{Load regulation} = V_{NL} - V_{FL}$$

where, V_{NL} = output voltage at no-load

V_{FL} = output voltage at full-load.

For instance, in the regulated power supply designed above, it has 12V across its output when the load is not connected, and, 11.8 V when the rated full load is connected. The load regulation of this power supply is given by,

$$\begin{aligned} \text{Load regulation} &= V_{NL} - V_{FL} \\ &= 12 \text{ V} - 11.8 \text{ V} = 0.2 \text{ V} \end{aligned}$$

Load regulation is generally expressed as a percentage. Percentage regulation is calculated as follows;

$$\% \text{ Load regulation} = \frac{V_{NL} - V_{FL}}{V_{NL}} \times 100\%$$

Percentage load regulation in the example considered above will be,

$$\% \text{ Load regulation} = \frac{12 - 11.8}{12} \times 100\% = 1.67\%$$

TIP: The lower the % load regulation value, the better is the quality of the power supply.

Line regulation

Line regulation is the change in the regulated load voltage for the specified range of input voltage or line voltage. Line voltage generally means the mains AC supply voltage.

For example, if the output voltage (also referred to as load voltage) changes from 12 V to 11.6 V when the line voltage changes from 240 to 200 V, then the line regulation of the power supply is,

$$\text{Line regulation} = 12 - 11.6 \text{ V} = 0.4 \text{ V}$$

Line regulation is also generally in percentage as,

$$\% \text{ Line regulation} = \frac{\text{Line regulation}}{\text{Nominal load voltage}} \times 100$$

For the above example,

$$\% \text{ Line regulation} = \frac{0.4}{12} \times 100 = 3.33\%$$

TIP: The lower the % line regulation the better is the quality of the regulated power supply.

Ripple rejection

Voltage regulators not only regulate or stabilize the output voltage against changes in the load current and input voltage but also reduce or attenuate the ripple that comes with the input voltage.

This is because of the fact that the ripple in the input is equivalent to a change in the input voltage, the voltage regulator reduces the effect of these changes at the output.

Ripple rejection (RR) is expressed as a ratio of the amount of ripple at the output and that at the input. RR is expressed usually in decibels. For example, if RR is specified as 80db, it means that the ripple at the output is 80 db or 10,000 times less than the ripple at the input.

With an intention to further reduce the ripple in the regulated output some series regulator circuits use a small value electrolytic capacitor at the output.

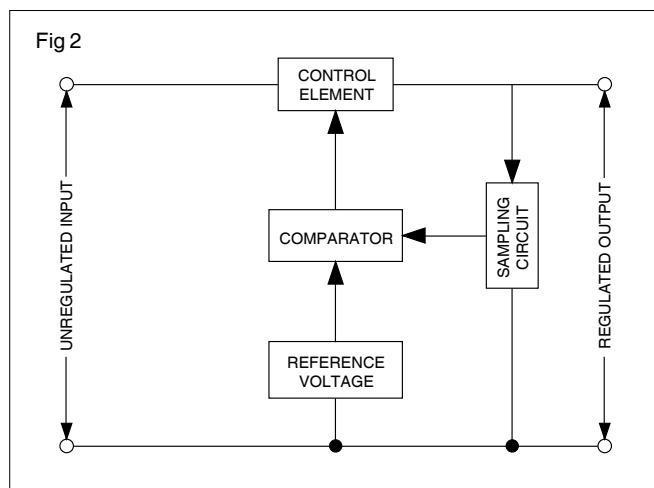
Summarising, the advantages of simple series regulators over zener regulators are,

- 1 reduced wattage requirement of zener diode
- 2 reduced output impedance
- 3 increased load and line regulation
- 4 reduced ripple at the output.

SERIES REGULATOR

A simple series regulator, consisting of a zener diode and an emitter follower was discussed. That was an improvement over the zener regulator. It is possible to further improve the voltage regulation of the simple series regulator by providing a voltage feed back as shown in Fig 3. To understand this circuit better, consider the block diagram of series voltage regulator with feed back as shown in Fig.2.

In Fig 2, the control element generally a pass transistor, controls the amount of input voltage that gets to the output. The output voltage is sampled by the sampling circuit. The sampled voltage provides a feed-back voltage to the comparator. The other input to the comparator is the reference voltage, generally using a zener diode.



If the output voltage increases (or tends to increase), the sampling circuit provides a sample of the increase to the comparator. The comparator circuit provides a control voltage (based on the amount of increased output) to the series element to decrease the output voltage, thereby maintaining the output voltage constant.

On the other hand, if the output voltage decreases (or tends to decrease) due to higher load current, the comparator provides a control signal to the series element to increase the output voltage, thereby maintaining the output voltage constant.

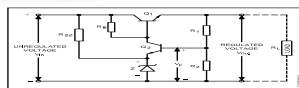
From Fig 2, it can be seen that the negative feedback provided will hold the output voltage almost constant despite the relative large changes in the input voltage and load current.

Practical transistor series regulator with feedback

Fig 3 shows a practical circuit schematic of a series regulator.

Transistor Q_1 is the pass transistor or the control element. As in a simple series regulator, the pass transistor Q_1 is connected as an emitter follower. Therefore, its base voltage is one V_{BE} (0.7V) higher than the output voltage V_{out} across the load.

Fig 3



Transistor Q_2 does the job of a comparator. This transistor Q_2 , operates in the active region as a linear amplifier.

The zener diode Z, provides the necessary reference voltage. In critical applications, zeners having a temperature coefficient approaching zero are used.

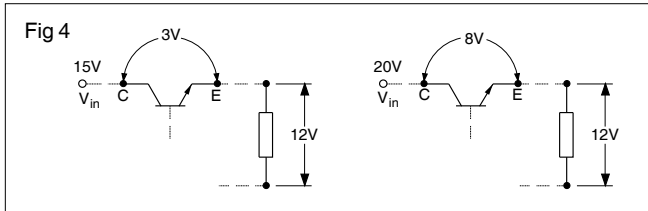
TIP: Zener voltages near 6 V have temperature a coefficient approaching zero.

The voltage divider network, R_1 and R_2 form the *sampling circuit*. The voltage divider samples the output voltage V_{out} and provides a feedback voltage V_F to the base of Q_2 .

The feedback voltage V_F varies the collector current of Q_2 . For instance, if V_{out} tries to increase, more V_F is fed to the base of Q_2 . This produces larger Q_2 collector current through R_B . Larger collector current of Q_2 results in increased voltage drop across R_B and reduced voltage at the collector of Q_2 . Reduced collector voltage at Q_2 means, less base voltage at Q_1 . The reduction in the base voltage of Q_1 results in reduced output voltage (recall Q_1 is an emitter follower).

Similarly, if the output voltage tries to decrease, there will be less feedback voltage V_F . Therefore, less base voltage at Q_2 resulting in less collector current of Q_2 . This results in more base voltage at Q_1 , and therefore, more output voltage across the load.

It is important to remember that the V_{CE} drop across the pass transistor Q_1 will be the difference between the input voltage V_{in} and the output voltage V_{out} as shown in Fig 4.



Simple design guidelines

Assuming an unregulated input of 34V to the regulator to be designed. Referring to Fig 3, the output voltage V_{out} is applied to the voltage divider formed by R_1 and R_2 . The feedback fraction k is given by,

$$k = \frac{R_2}{R_1 + R_2} \quad \dots\dots\dots[1]$$

TIP: The feedback factor k will always be less than 1.

The feedback voltage V_F will be,

$$V_F = kV_{out}$$

$k.V_{out}$ is also given by,

$$k.V_{out} = V_Z + V_{BE}$$

Therefore,

$$V_{out} = \frac{V_Z + V_{BE}}{k} \quad \dots\dots\dots[2]$$

where,

V_{out} = Regulated output voltage

k = Feedback factor

V_Z = Zener voltage

V_{BE} = base-emitter voltage of Q_2 .

If $R_1 = 330\Omega$ and $R_2 = 1.5 K\Omega$, using eqn 1,

$$k = \frac{1500}{330 + 1500} = 0.819.$$

If a 9.1V zener is chosen, then V_{out} will be, using eqn 2,

$$V_{out} = \frac{9.1 + 0.7}{0.819} = 11.965 \approx 12 \text{ volts.}$$

Transistor Q_2 can be any silicon small signal transistor such as BC 147.

Current through the voltage divider R_1, R_2 will be,

$$I_d = \frac{V_{out}}{R_1 + R_2} = \frac{12 \text{ V}}{1830 \Omega} = 6.5 \text{ mA.}$$

TIP: I_d should be preferably below 10mA. This is because this current is also a load to the pass transistor.

Assuming a maximum load current I_L of 60 mA, the pass transistor has to pass this current. Assuming the worst case β of 80 for the pass transistor,

$$I_B \text{ of } Q_1 = \frac{I_L + I_d}{\beta} = \frac{60\text{mA} + 6.5\text{mA}}{80} = 0.83\text{mA} \text{ or } 830 \mu\text{A.}$$

The base current of Q_1 has to be supplied by Q_2 . If $R_B = 3.3K$, the voltage at the collector of Q_2 will be one V_{BE} drop more than the V_{out} (12V). That is,

$$V_C \text{ of } Q_2 = 12 + 0.7 = 12.7.$$

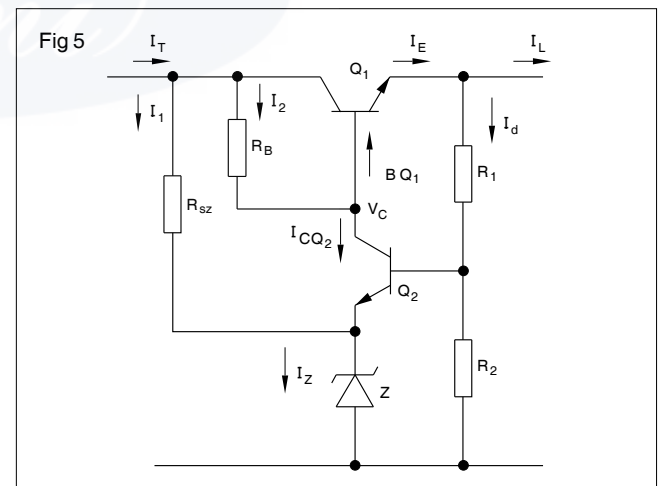
Therefore, current I_1 through R_B will be,

$$I_2 = \frac{V_{in} - V_C \text{ at } Q_2}{R_B} = \frac{18\text{V} - 12.7\text{V}}{3300\Omega} = 1.6\text{mA.}$$

NOTE: Unregulated input voltage V_{in} given at the time of design is 18V.

As shown in Fig 5, collector current of Q_2 is,

$$I_{CQ2} = I_2 - I_{BQ1} = 1.6\text{mA} - 0.83\text{mA} = 0.77\text{mA.}$$



If R_{SZ} is 820 Ω , current I_1 through it is,

$$I_1 = \frac{18 - 9.1}{820\Omega} \approx 10.85\text{mA.}$$

Current I_Z through the zener diode is,

$$I_Z = I_1 + I_{EQ2} = 10.85\text{mA} + 0.77\text{mA} = 11.62\text{mA.}$$

$I_{EQ2} \approx I_{CQ2}$. Since $I_{CQ2} = 0.77\text{mA}$, I_{EQ2} is also taken as 0.77mA.

Zener wattage rating is given by,

$$P_Z = V_Z \times I_Z = 9.1\text{V} \times 11.62\text{mA} \approx 106 \text{ m watts.}$$

Hence a 1/4W zener is sufficient for this circuit.

Maximum current, I_E through the pass transistor Q_1 is, the sum of, load current I_L and voltage divider circuit current I_d which is equal to $60\text{mA} + 6.5\text{mA} = 66.5\text{mA}$.

Therefore, Power dissipation across Q_1 at maximum input voltage and maximum load current is,

$$P_D = V_{CE} \times I_E = (18 - 12)(66.5\text{mA}) = 399 \text{ m watts.}$$

Any medium power silicon transistor which has a power rating of atleast 500 m watt, such as SL 100 or CD100 can be chosen as the Pass transistor.

Summarizing, the values of components in the feed-back series regulator.

Circuit shown in Fig 3 are,

Q_1 : Transistor with $\beta(\text{min})$ of 80

Power dissipation 500mw(min)

Collector current 66.5mA(min)

Q_2 : Transistor with $\beta(\text{typ})$ of 80

Collector current 0.77mA(min)

Power dissipation 25mw(min)

Z : 9.1V zener, 1/4watt

R_{SZ} : 1.8KW, 1/2 Ω , 10%

R_B : 3.3KW, 1/4 Ω , 10%

R_1 : 330W, 1/4 Ω , 5%

R_2 : 1.5KW, 1/4 Ω , 5%

More accurate the values of R_1 and R_2 , better will the feedback voltage and hence the regulation. Therefore the tolerance value of R_1 and R_2 is chosen as 5%.



Fixed & variable regulators using 3 pin ICs

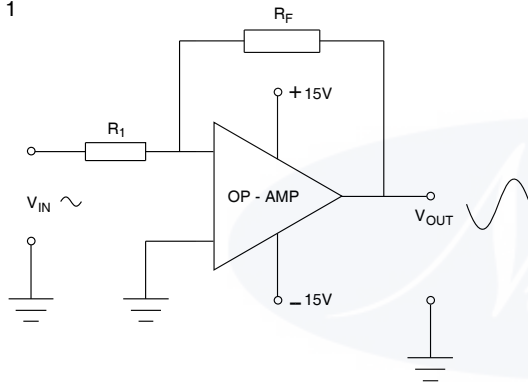
Objectives: At the end of this lesson you shall be able to

- draw the circuit diagram of a dual power supply
- list a few variable regulator 3-pin ICs
- draw the circuit diagram of a variable output regulator using a LM317 IC
- draw the circuit diagram of a variable output regulator using a 723 IC
- state the functions of the external components in regulator circuits.

Most electronic circuits generally need either a +ve DC supply or a -ve DC supply for its working. However, there are circuits which are designed to work using both + and - supplies. An example of circuits which require both +ve and -ve supply are the OP-AMPs. OP-AMPs are integrated circuit amplifiers which need, +ve supply, -ve supply and ground. A typical OP-AMP circuit is shown in Fig 1.

Details of OP-AMPs and its application is discussed in further lessons.

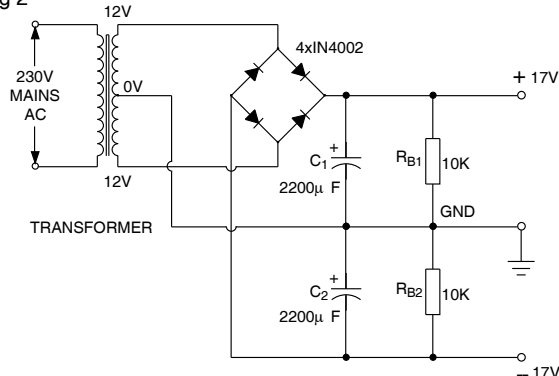
Fig 1



Therefore, for circuits which require both + and - DC supplies, a single power supply which can deliver both $\pm$ DC is required to be designed. Power supplies which can deliver both $\pm$ DC are generally referred to as *Dual Power Supply*.

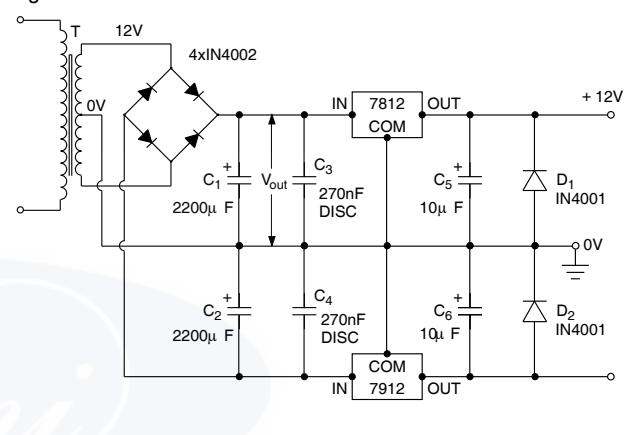
To design a $\pm$ or dual regulated power supply, as a first step it is required to design a $\pm$ unregulated DC supply. Fig 2 shows a simple method of obtaining $\pm$ unregulated DC supply.

Fig 2



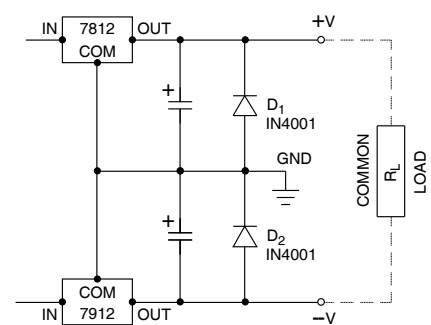
Once, a $\pm$ unregulated DC supply is available, one each of +ve regulator 3-terminal IC and a -ve regulator 3-terminal IC can be attached, to obtain a $\pm$ regulated DC supply. One such $\pm$ regulated DC supply using 7812 (+ve regulator) and 7912 (-ve regulator) is shown in Fig 3.

Fig 3



For +ve and -ve regulator circuits shown in Fig 3 used the diodes D_1 and D_2 . The function of these diodes is very important. If these diodes are not used, the regulator ICs may get damaged due to *common load* problems. The term common load means, a load connected across the +ve and -ve outputs of the regulator as shown in Fig 4. Because of the fact that these common leads does not make use of the ground (GND) several problems occur when the supply is switch ON, in case of over loads and so on (For further details on common-load problems in dual supplies refer reference books listed at the end of this book). Hence to avoid the common load problem in dual power supplies diodes D_1 and D_2 are very essential.

Fig 4



Variable/adjustable output voltage regulators

A number of IC voltage regulators are available using which an adjustable output voltage of 1.2V to 32 volts can be obtained. Amongst these adjustable output voltage regulators, there are two types:

- 3-Terminal variable output voltage regulators ICs
- Multi-terminal variable output voltage regulator ICs

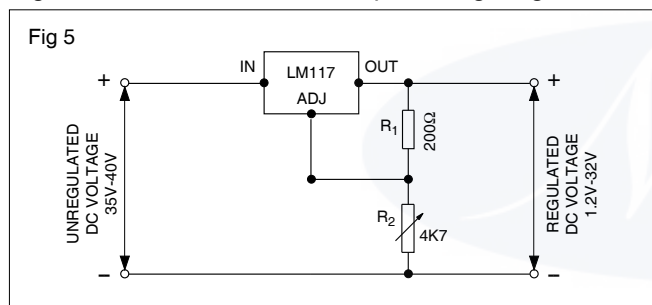
3-Terminal variable output regulators ICs

These ICs look like 3-terminal fixed output voltage regulators as shown in Fig 3. A few examples of 3-terminal adjusted output voltage regulator ICs are,

- LM117 Output adjustable from 1.2 V to 37 V
- LM317 Output adjustable from 1.2 V to 32 V
- LM338 Output adjustable from 1.2 V to 32 V
- LM350 Output adjustable from 1.2 V to 33 V

These variable output voltage regulator ICs are designed for adjustable output voltage, unlike the fixed output 3-pin regulators such as 7812, LM 340-5 etc which can be modified to get variable output voltage.

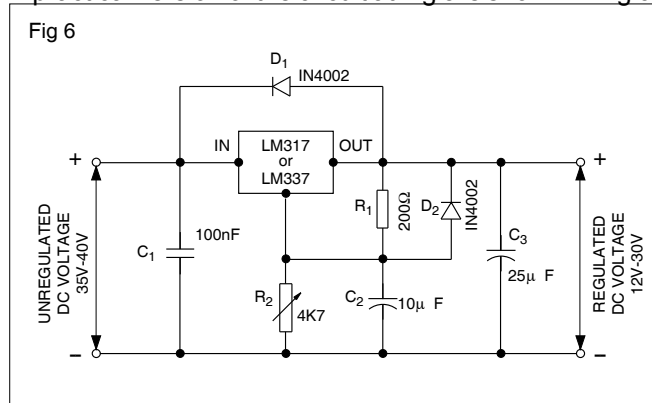
Fig 5 shows a basic variable output voltage regulator.



In the circuit at Fig 3, if the adjustment terminal (ADJ) is grounded, the output of the regulator will be 1.2 volts. To obtain a higher output voltage a small reference voltage is given at ADJ using a voltage divider circuit consisting of R_1 and R_2 as shown in Fig 5. With this the regulated output voltage is approximately given by

$$V_{out} = 1.2V \times \left(1 + \frac{R_2}{R_1} \right) \dots\dots\{1\}$$

A practical version of the circuit at Fig 3 is shown in Fig 6.



This circuit uses a few bypass capacitors and protection diodes.

In Fig 6, capacitor C_1 is used to prevent setting up of the oscillations and should be connected as close to IC as possible. Capacitor C_2 is used to improve the ripple in the output voltage. Note that the value of C_3 should not be very high (recall, surge current). Capacitor C_2 is used to avoid excess ringing.

When external capacitors are used with any IC regulator, it is necessary to add protection diodes to prevent the capacitors from discharging through low current points into the regulator. Hence, diodes D_1 and D_2 are used. D_1 protects the IC against shorts due to C_3 and D_2 protects against shorts due to C_2 .

The ICs LM317 and 338 have built in fold back current limiting and thermal protection. These IC are available in both plastic and metal packages with current ratings from 0.1A (LM317L) to 5A(LM338K). For other details refer Appendix.

LM117, LM317 and LM338 are of the same family ICs, and hence, are interchangeable. For details refer the National Semiconductor application notes.

Multiple-pin-Variable voltage regulator ICs

Unlike 3-pin fixed output voltage regulators and 3-pin variable output voltage regulators, voltage regulator ICs having multiple pins are designed for versatility. These multiple pin IC regulators can be used as a linear regulator (all the regulators discussed so far), or as a switching regulator (to be discussed), or as a shunt regulator (to be discussed) or as a current regulator (to be discussed).

Generally in multiple pin types of regulators, dissipation limitation of the IC packages restrict the output current to a few tens of milliamps. However, external transistors can be added to obtain currents in excess of 5A.

Some of the multiple pin, versatile IC regulators are, LM100, LM105, LM205, LM305, μ A723, CA3085 and so on.

Practical circuits using 723 IC

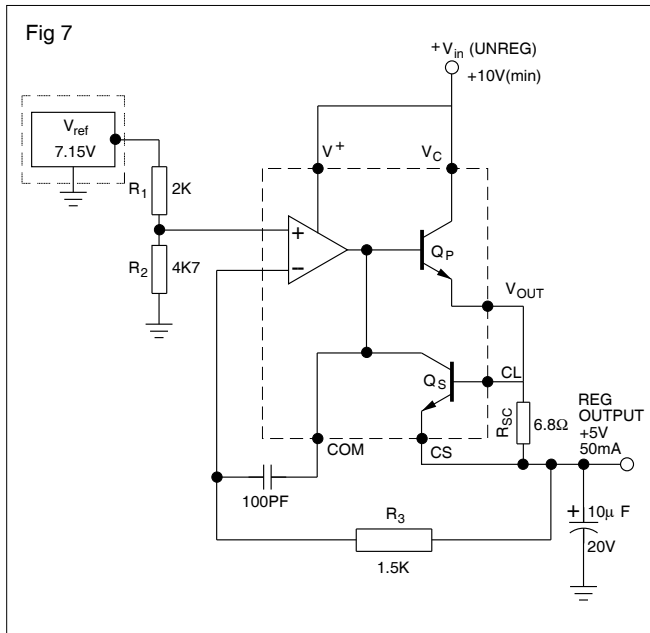
Fig 7 shows a +5V, 50mA regulator using 723 IC.

In Fig 5, the internal components of the regulator IC are shown in dotted lines. The IC has a internal reference voltage of 7.15 V. This reference voltage can be varied using a voltage divider R_1, R_2 . In Fig 7, the regulated output voltage is given by,

$$V_{out} = V_{ref} \frac{R_2}{R_1 + R_2} = 7.15V \left(\frac{4K7}{2K + 4K7} \right) = 5V.$$

The value of R_3 in Fig 7 should be chosen such that there is minimal temperature drift. The value of R_3 should be,

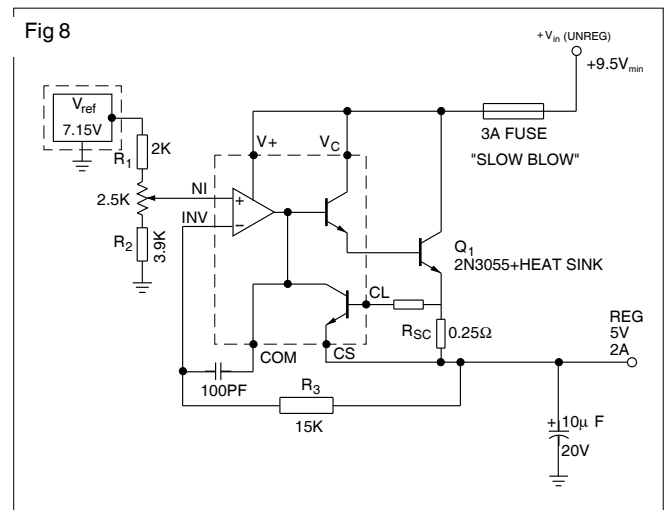
$$R_3 = \frac{R_1 R_2}{R_1 + R_2}$$



Resistor R_{SC} is the short circuit current sensing resistor. The value of R_{SC} is chosen for about 0.5 volts drop at the maximum desired output current. This 0.5V across CL-CS input will turn on the current limiting transistor, Q_S shutting off the base drive to the output pass transistor Q_P .

Higher output current using external transistor

Fig 8 shows a +5V, 2 amps regulator which is a modified form of the regulator shown in Fig 7.



In most cases, the power supply requirement is satisfied by the newer 3-pin variable regulators. Hence, multiple-pin IC regulators are rarely used.

Troubleshooting simple power supplies

Objectives: At the end of this lesson you shall be able to

- list the activities involved in troubleshooting
- list the three general steps involved in troubleshooting
- list and explain the two popular methods of troubleshooting
- list the possible defects in a simple power supply with filter
- state the meaning and use of Problem trees(PT)
- state the meaning and use of Service-flow-diagrams(SFD)

INTRODUCTION

Troubleshooting any equipment or a circuit involves the following activities:

- To identify the exact nature of the problem.
- To identify the section causing the problem.
- To isolate and arrive at the exact cause(s).
- To confirm the causes by necessary tests.
- To replace the problem-causing parts.
- To re-test and confirm the satisfactory working.

The following are the general steps involved in troubleshooting:

1 Physical and sensory tests

- Look for the most common physical faults, such as broken wires, cracked circuit boards, dry solders and burn components.

- Smell for hot or burning components.

- Feel with the fingers for unduly hot components.

2 Symptom diagnosis

Learn the operation of the system to be repaired with the help of its block diagram and its input and output specifications.

Observe the symptoms produced by the defective system, and determine which section or function would produce the symptoms.

3 Testing and replacing defective components

When the probable defective section has been diagnosed, check the probable components in that section of the circuit that are most likely to go defective in the order given below:

Components should be checked in the order given below because that is the order in which they fail in most cases.

(a) Active high power components

For example, components such as transistors, ICs, and diodes. High power devices are physically large in size and are used for handling the high power, generally in output circuits.

(b) Active low power components

These are the same as in (a) but are physically small and can handle smaller amounts of power.

(c) High voltage/power passive components

Such components are resistors, capacitors, transformers, coils, etc. which handle large amounts of voltage/power. They are found in power supplies and output circuits.

(d) Low power passive components

These are the same as in (c) but are physically smaller and handle comparatively less power and are low in value (ohms, microfarads, microhenries, etc.)

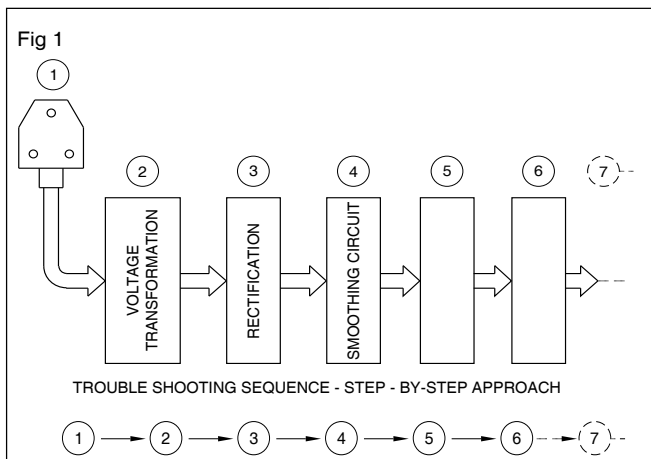
This procedure may not turn out to be true always. Hence, do not attempt to replace common sense and meter measurements with the procedure.

While troubleshooting any electronic system, two main methods are generally adopted. They are:

Step-by-step method of troubleshooting

This approach is preferred by the beginners. In this approach, the problem causing part or section is identified by testing the parts or sections from the beginning to the end as shown in Fig 1.

Although this approach may take more time, this is the most suited approach for the beginners.



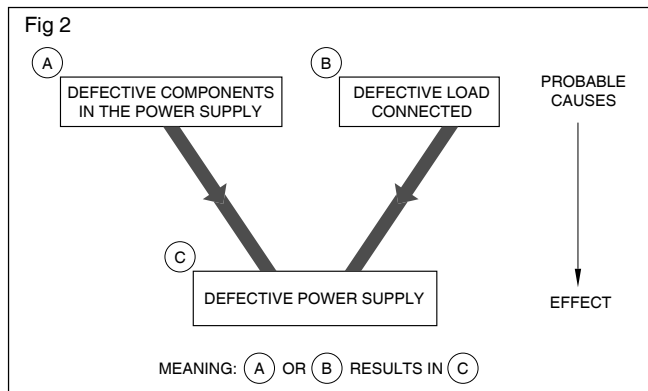
Short cut or logical approach method of troubleshooting

This method is used by the experienced servicing people. In this method, the problem causing part or section is identified from the nature of the problem symptom. Divide and conquer procedure is adopted to arrive at the exact cause. This method takes less time comparatively.

Troubleshooting Power supplies

All electronic systems can be broken down into blocks, generally based on their function. Fig 1 shows the various blocks of a simple power supply. Each block has a particular function to perform.

Before carrying out the troubleshooting of power supplies, the first thing to be done is to isolate the load connected to the power supply. This is because the connected load itself may be the cause of the problem as shown in the problem tree (PT) in Fig 2.



Once it is confirmed that the power supply has the same defect even with the load disconnected, you can follow either the step-by-step approach or the logical approach to troubleshoot the power supply.

STEP-BY-STEP approach to troubleshoot power supply

In the step-by-step approach of troubleshooting, the various blocks of the power supply shown in Fig 1 and the components of the blocks are checked one by one, starting with block 1 and in steps as given below.

Step 1: Confirm the presence and satisfactory level of the mains supply from which the power supply is powered.

Step 2: Switch the power ON and test and note down the exact nature of the problem. Although the nature of the problem has been already told, it is essential to confirm the exact nature of the problem. This is because, in a real life situation, the customer may not be a technical person to inform the exact nature of the problem.

Step 3: Carry out physical and sensory tests.

Step 4: Trace the circuit to identify any wrong polarity connections.

Step 5: Remove the power cord of the power supply from the mains and test the power cord.

Step 6: Test the transformer.

Step 7: Test the diode(s) of the rectifier section.

Step 8: Test the capacitor(s) of the filter section.

Step 9: Test the bleeder resistor, surge resistor and other resistors, if any.

Step 10: Test the output indicator lamps/LEDs.

After completing all the above steps, from the defective components identified, analyze the root cause for the problem and confirm that the cause will not reoccur if the identified components are replaced.

Step 11: Replace the identified defective component(s).

Step 12: Switch the power ON and test the power supply, first without load, and then connecting it to the load.

LOGICAL approach to troubleshoot power supply

In this approach steps 1 to 4 of the step-by-step approach are the same. The next step is to refer to the Logical Service Flow Diagram(SFD) for the identified problem and proceed with the troubleshooting as directed in the SFD.

SFDs are very good tools in troubleshooting as they take into account, the divide and conquer technique, thus reducing the overall time taken to troubleshoot the defect in the power supply.

The possible types of defects that can occur in a simple power supply consisting of a bridge rectifier and capacitance input filter are listed below along with their SFD numbers.

Possible defects in a power supply using Bridge rectifier and filter capacitor

[1] No output voltage

This defect in the power supply may be due to one or more component of the circuit. Problem Tree-1(PT-1) given at the end of this lesson for the causes of the problem No output voltage.

This PT shows the cause-effect relationship of the defective components with the problem. The cause is given at the top and the effect at the bottom for the only reason that it is a normal tendency to read a page from the top to the bottom.

PT-1 shows two problem trees. The first tree in Chart 1, indicated as Level-1 is a simple tree which gives the level-1 causes of the problem. Level-2 is an extension of the same problem tree, which gives one more level of the causes for the causes given in the simple tree at level-1.

[H.I: Instructor to discuss PT-1 and ensure that trainees clearly understand the need and meaning of PTs.]

Chart 2 at the end of this lesson, shows the sequence to be adopted while servicing a defective power supply. The Service Flow Sequence -1(SFS-1) at Chart 2 is self-explanatory. However the following tips make it easy to go through the SFS.

- 1] The flow is from Top to Bottom.
- 2] Rectangular blocks indicate work to be done or action to be taken.
- 3] Follow the path of the arrow.
- 4] Diamond blocks indicate a decision to be taken after conducting a test or making a measurement. If the answer to the question in the diamond block is YES, follow the path of YES and continue. If the answer is NO, follow the path of NO and continue.
- 5] Rounded rectangular block indicates the end of the job.

[2] Low output voltage/ Increased ripple in output

Here, note that two defects are combined. The reason being, that these two defects generally occur simultaneously. If the output voltage is low, the cause(s) for this also results in increased ripple and vice versa. Ofcourse with one exception that, if the mains level itself is low or if the secondary voltage of the transformer itself is low due to shorted windings, a low output voltage is not associated with increased ripple.

The causes for this defect are given in problem tree PT-2 in Chart 3. Chart 4 shows, the service flow diagram(SFS-2) for servicing the defect.

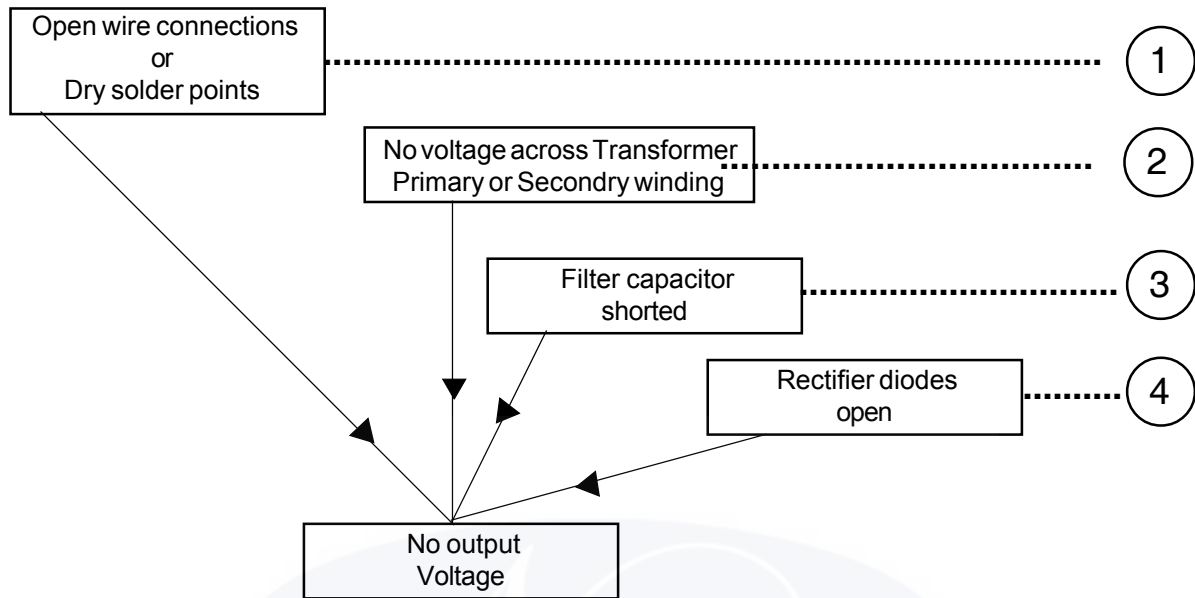
The SFSs and PTs for a fullwave rectifier with a capacitance filter is almost similar to that of a bridge rectifier. However, it is suggested that the trainees shall make SFSs and PTs for a fullwave rectifier power supply on their own for practice and better understanding of the method.

Problem Tree - PT1

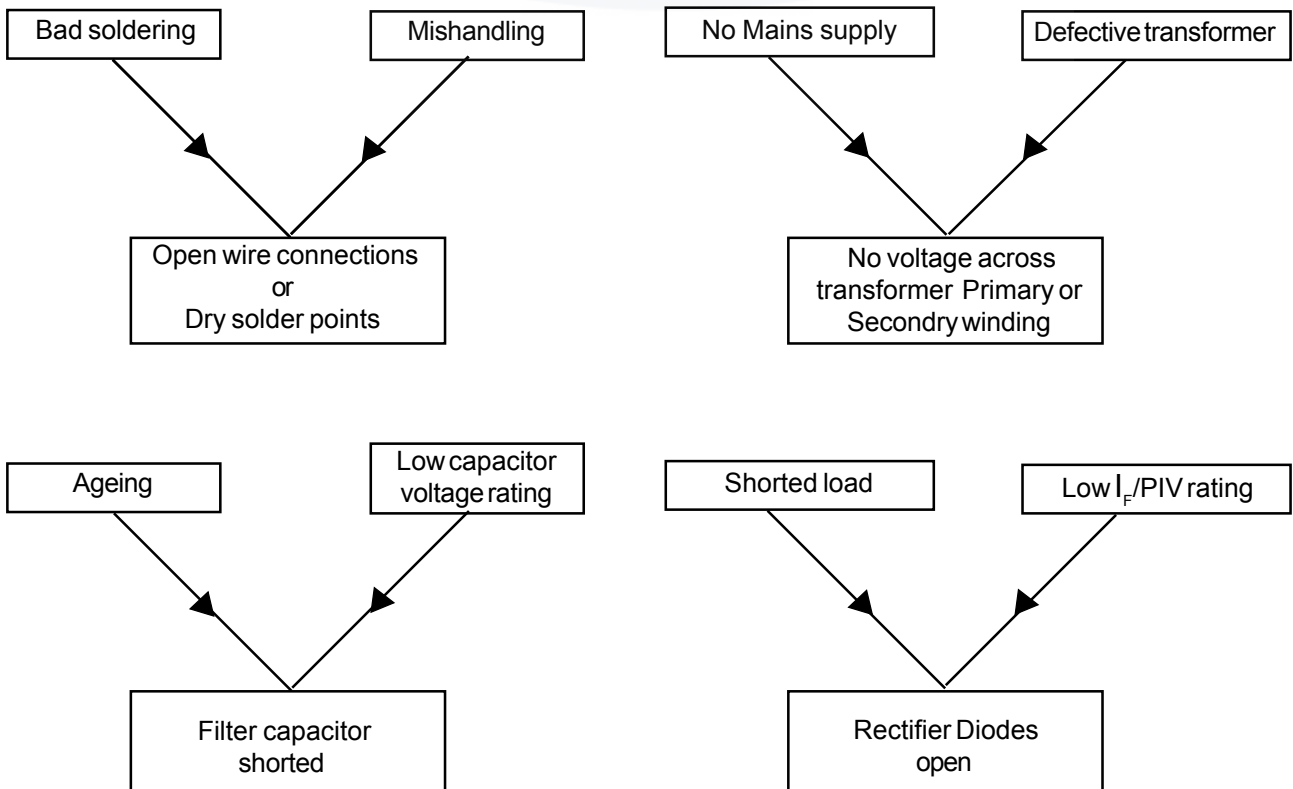
NATURE OF DEFECT : No Output voltage
 TYPE OF SYSTEM : Bridge rectifier with capacitor filter

LEVEL 1

Possibility
order

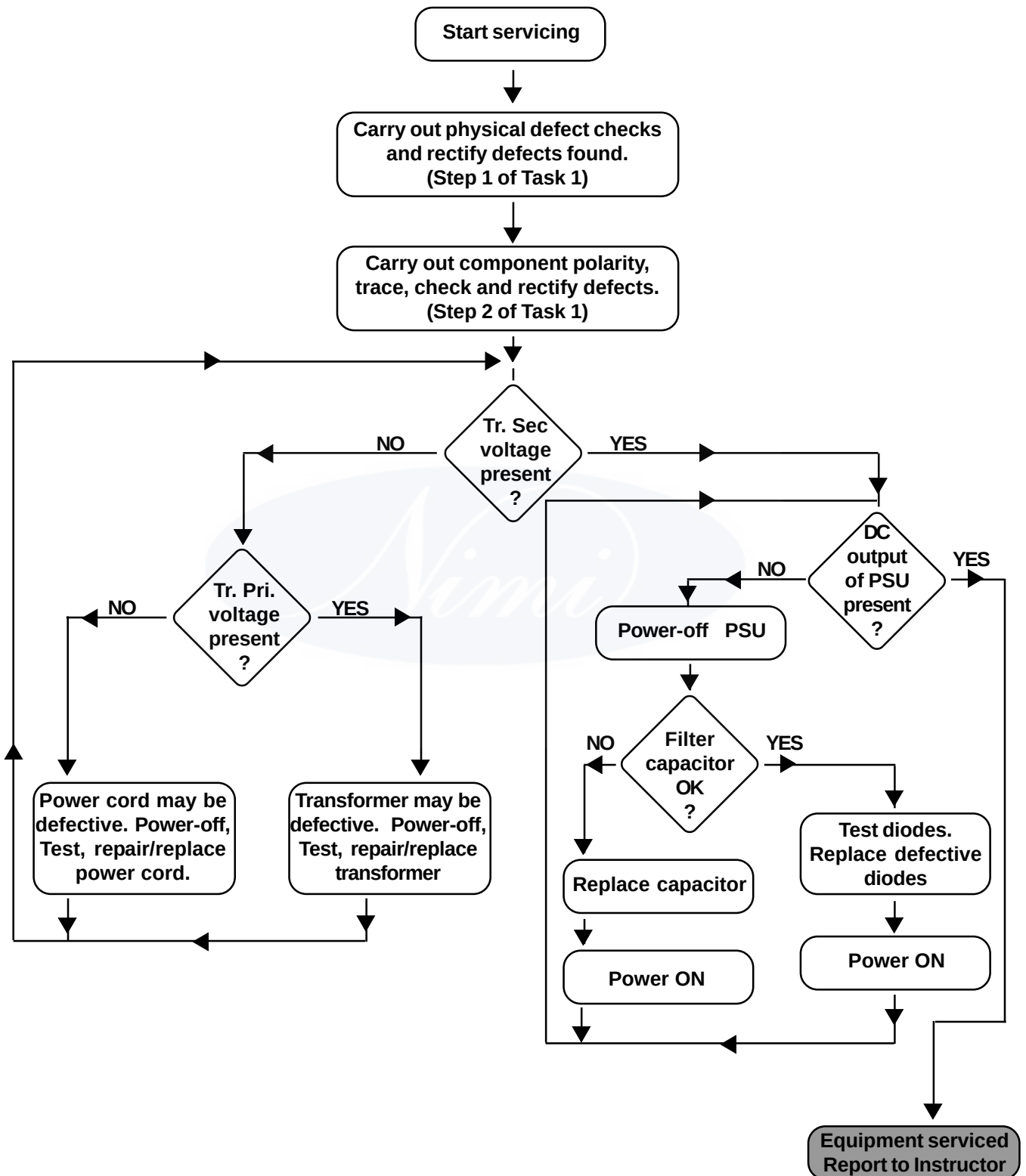


LEVEL 2



Service flow sequence (SFS-1)

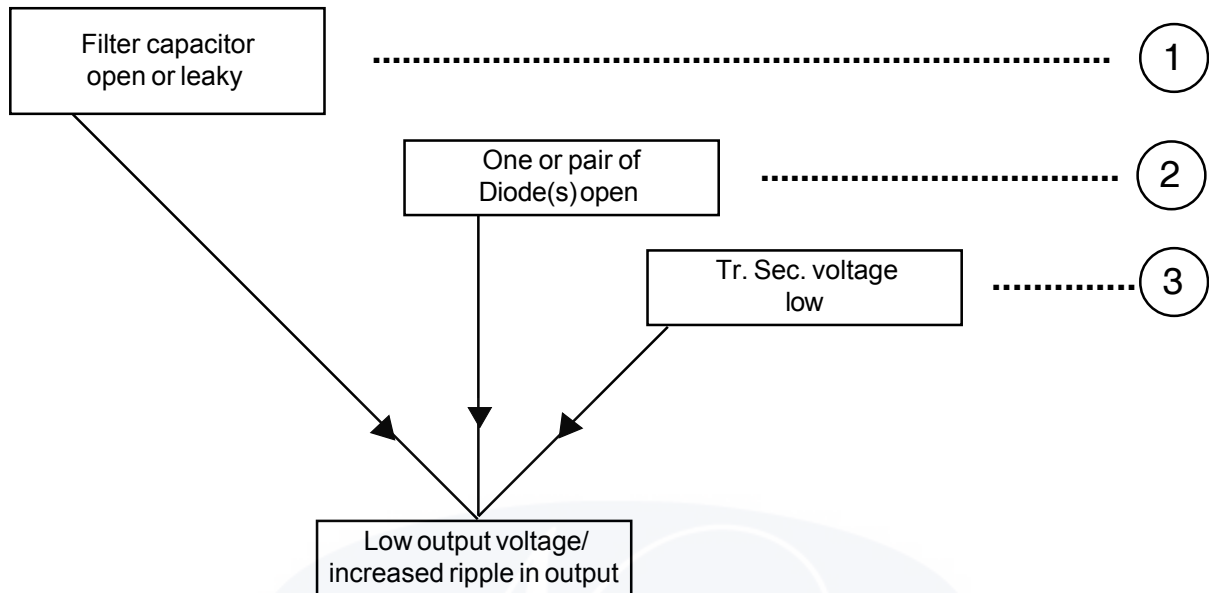
NATURE OF DEFECT : Defective power supply with NO OUTPUT VOLTAGE



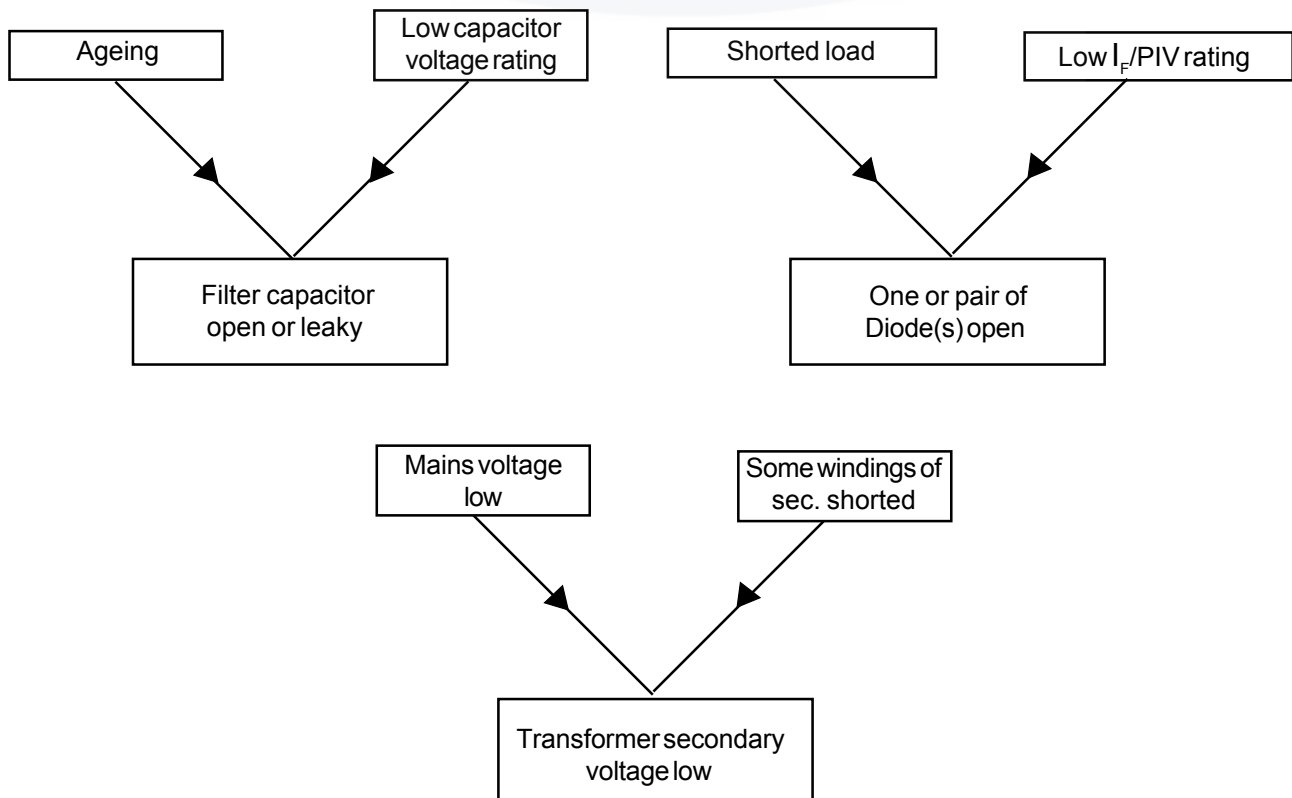
Problem Tree - PT2

NATURE OF DEFECT : Low Output DC/Increased ripple
 TYPE OF SYSTEM : Bridge rectifier with capacitor filter

LEVEL 1

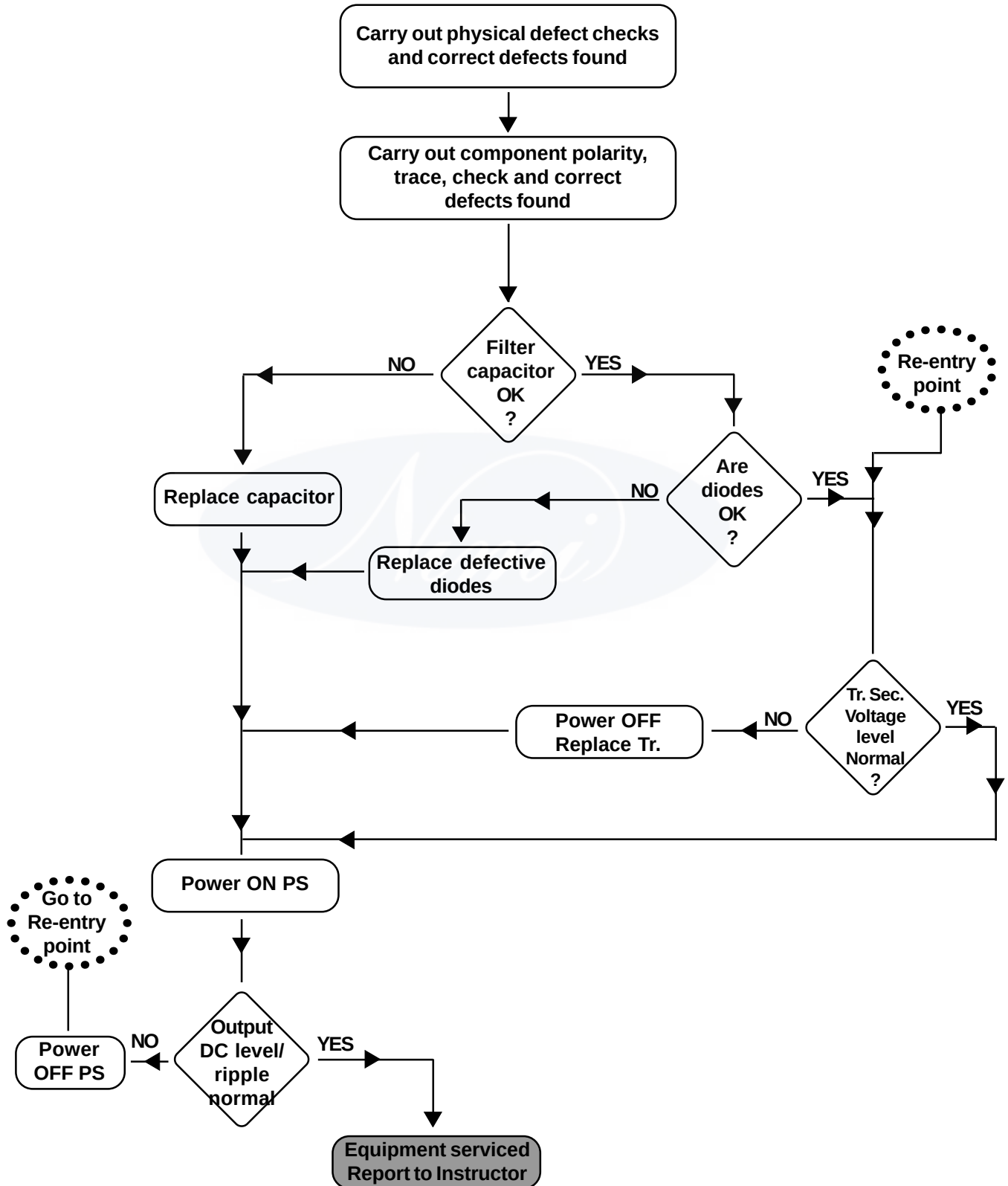


LEVEL 2



Service flow Sequence (SFS-2)

NATURE OF DEFECT : Low output DC/ Increased ripple



Working of inverters

Objectives : At the end of this lesson you shall be able to

- define the detail study of inverter
- design inverter
- define application of inverter.

An **inverter** is an electronic circuit that converts direct current (DC) to alternating current (AC). Inverters are used in a wide range of applications, from small switching power supplies in computers, to large electric utility applications that transport bulk power. The inverter is so named because it performs the opposite function of a rectifier.

Inverter applications

The following are examples of inverter applications.

Fig. 1



MEL1101

DC power source utilization

Inverter designed to provide 115 VAC from the 12 VDC source provided in an automobile

An inverter allows the 12 or 24 volt (battery) DC power available in an automobile or from solar panels to supply AC power to operate equipment that is normally supplied from a main power source.

Inverters are also used to provide a source of AC power from photovoltaic solar cells and fuel cell power supplies.

Uninterruptible power supplies

One type of uninterruptible power supply uses batteries to store power and an inverter to supply AC power from the batteries when main power is not available. When main power is restored, a rectifier is used to supply DC power to recharge the batteries.

Induction heating

Inverters convert low frequency main AC power to a higher frequency for use in induction heating. To do this, AC power is first rectified to provide DC power. The inverter then changes the DC power to high frequency AC power.

High-voltage direct current (HVDC) power transmission

With HVDC power transmission, AC power is rectified and high voltage DC power is transmitted to another location. At the receiving location, an inverter in a static inverter plant converts the power back to AC.

Variable-frequency drives

A variable-frequency drive controls the operating speed of an AC motor by controlling the frequency and voltage of the power supplied to the motor. An inverter provides the controlled power. In most cases, the variable-frequency drive includes a rectifier so that DC power for the inverter can be provided from main AC power. Since an inverter is the key component, variable-frequency drives are sometimes called inverter drives or just inverters.

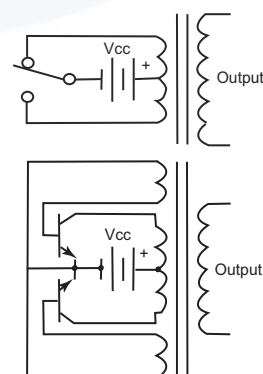
Electric vehicle drives

Adjustable speed motor control inverters are currently used in some electric locomotives and diesel-electric locomotives as well as some battery electric vehicles and hybrid electric highway vehicles such as the Toyota Prius. Various improvements in inverter technology are being developed specifically for electric vehicle applications.

Inverter circuit description

Simple inverter circuit shown with an electromechanical switch and with a transistor switch is shown in Fig 2.

Fig 2



MEL1102

Basic inverter designs

In one simple inverter circuit, DC power is connected to a transformer through the centre tap of the primary winding. A switch is rapidly switched back and forth to allow current to flow back to the DC source following two alternate paths through one end of the primary winding and then the other. The alternation of the direction of current in the primary winding of the transformer produces alternating current (AC) in the secondary circuit.

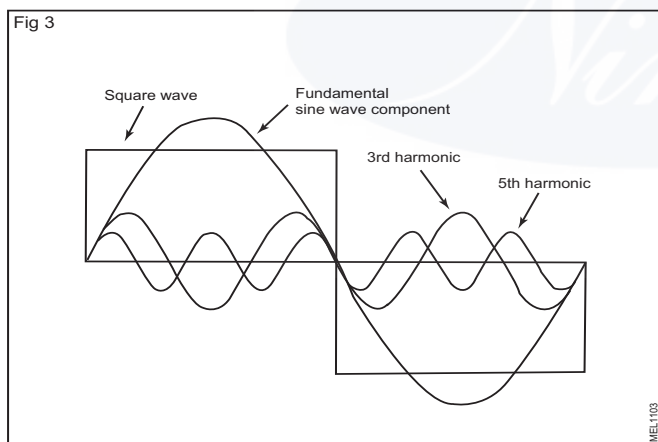
The electromechanical version of the switching device includes two stationary contacts and a spring supported moving contact. The spring holds the movable contact against one of the stationary contacts and an electromagnet pulls the movable contact to the opposite stationary contact. The current in the electromagnet is interrupted by the action of the switch so that the switch continually switches rapidly back and forth. This type of electromechanical inverter switch, called a vibrator or buzzer, was once used in vacuum tube automobile radios. A similar mechanism has been used in door bells, buzzers and tattoo guns.

As they have become available, transistors and various other types of semiconductor switches have been incorporated into inverter circuit designs.

Square waveform with fundamental sine wave component, 3rd harmonic and 5th harmonic

Inverter output waveforms

The switch in the simple inverter described above produces a square voltage waveform as opposed to the sinusoidal waveform that is the usual waveform of an AC power supply. Using Fourier analysis, periodic waveforms are represented as the sum of an infinite series of sine waves. The sine wave that has the same frequency as the original waveform is called the fundamental component. The other sine waves, called harmonics, that are included in the series have frequencies that are integral multiples of the fundamental frequency.



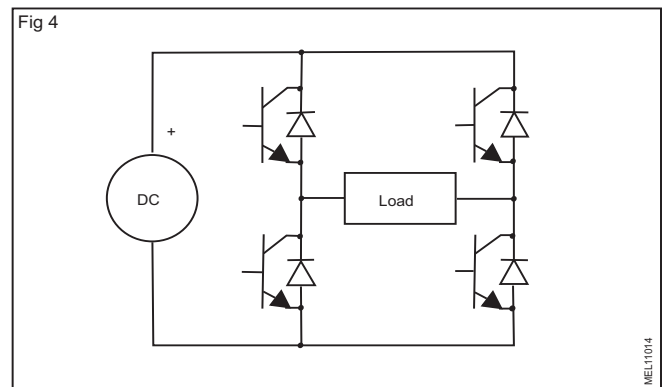
The quality of the inverter output waveform can be expressed by using the Fourier analysis data to calculate the total harmonic distortion (THD). The total harmonic distortion is the square root of the sum of the squares of the harmonic voltages divided by the fundamental voltage:

$$\text{THD} = \frac{\sqrt{V_2^2 + V_3^2 + V_4^2 + \dots + V_n^2}}{V_1}$$

The quality of output waveform that is needed from an inverter depends on the characteristics of the connected load. Some loads need a nearly perfect sine wave voltage supply in order to work properly. Other loads may work quite well with a square wave voltage.

More advanced inverter designs

H-Bridge inverter circuit with transistor switches and antiparallel diodes are shown in Fig 4.



There are many different power circuit topologies and control strategies used in inverter designs. Different design approaches address various issues that may be more or less important depending on the way that the inverter is intended to be used.

The issue of waveform quality can be addressed in many ways. Capacitors and inductors can be used to filter the waveform. If the design includes a transformer, filtering can be applied to the primary or the secondary side of the transformer or to both sides. Low-pass filters are applied to allow the fundamental component of the waveform to pass to the output while limiting the passage of the harmonic components. If the inverter is designed to provide power at a fixed frequency, a resonant filter can be used. For an adjustable frequency inverter, the filter must be tuned to a frequency that is above the maximum fundamental frequency.

Since most loads contain inductance, feedback rectifiers or antiparallel diodes are often connected across each semiconductor switch to provide a path for the peak inductive load current when the semiconductor is turned off. The antiparallel diodes are somewhat similar to the freewheeling diodes used in AC/DC converter circuits.

Fourier analysis reveals that a waveform, like a square wave, that is antisymmetrical about the 180 degree point contains only odd harmonics, the 3rd, 5th, 7th etc. Waveforms that have steps of certain widths and heights eliminate or "cancel" additional harmonics. For example, by inserting a zero-voltage step between the positive and negative sections of the square-wave, all of the harmonics that are divisible by three can be eliminated. That leaves only the 5th, 7th, 11th, 13th etc. The required width of the steps is one third of the period for each of the positive and negative voltage steps and one sixth of the period for each of the zero-voltage steps.

Changing the square wave as described above is an example of pulse-width modulation (PWM). Modulating, or regulating the width of a square-wave pulse is often used as a method of regulating or adjusting an inverter's output voltage. When voltage control is not required, a fixed pulse width can be selected to reduce or eliminate selected harmonics. Harmonic elimination techniques are generally applied to the lowest harmonics because filtering is more effective at high frequencies than at low frequencies. Multiple pulse-width or carrier based PWM control schemes produce waveforms that are composed of many narrow pulses. The frequency represented by the number of narrow pulses per second is called the switching frequency or carrier frequency. These control schemes are often used in variable-frequency motor control inverters because they allow a wide range of output voltage and frequency adjustment while also improving the quality of the waveform.

Multilevel inverters provide another approach to harmonic cancellation. Multilevel inverters provide an output waveform that exhibits multiple steps at several voltage levels. For example, it is possible to produce a more sinusoidal wave by having split-rail direct current inputs at two voltages, or positive and negative inputs with a central ground. By connecting the inverter output terminals in sequence between the positive rail and ground, the positive rail and the negative rail, the ground rail and the negative rail, then both to the ground rail, a stepped waveform is generated at the inverter output. This is an example of a three level inverter: the two voltages and ground.

Introduction to UPS

Objectives : At the end of this lesson you shall be able to

- **state the purpose of an UPS**
 - **state the types of UPS system**
 - **list the advantages and disadvantages of an off-line UPS**
 - **describe different front panel indications, rear panel sockets and switches**
 - **state the type of protections used in UPS**
 - **list the tips for testing the UPS**
 - **define the types of change over in off-line system.**
-

Basics of UPS systems

For ordinary house hold appliances such as incandescent lamps , tubes, fans, TV and fridge, the main AC supply doesn't make much of a difference. But when used for computers, medical equipments, telecommunication systems , a clean stable, interruption-free power supply is of utmost importance.

As more and more personnel computers word processors and data terminals find their way into small business, UPS systems that meet the power requirement and power range needs for small business and offices are being manufactured.

Earlier operating systems (DOS) doesn't have any shutdown procedure. So in case power failure it doesn't affect the operating system. Latest operating system Windows 9x and application softwares require proper shutting down and exit procedures. This procedures requires time which is provided by UPS in case of mains power failure.

UPS (Uninterrupted Power Supply) is the only solution available to an individual customer faced with the problem of ensuring high quality of power for critical loads. All UPS designs contain a battery charger to keep the battery fully charged by the power from mains. Small UPS normally comes with a sealed maintenance free (SMF) batteries which can provide 10 to 15 minutes of power backup, the backup time increases with the capacity of the battery. Tubular batteries or automotive batteries are used in medium and large capacity UPSs.

UPS classification

There are two broad categories of UPS topologies - OFF line, and ON line. These topologies differ in the way they serve the load when the mains is present and is healthy.

They vary in features & pricing.

Offline and online

Offline UPS filters the mains and feeds it directly to the load for most of the time. When the mains is unhealthy, perhaps due to a slight dip in voltage, the load is switched by a fast relay, in typically less than half a cycle, to an inverter deriving its power from a battery. The inverter generates a square or stepped waveform to emulate the mains-satisfactorily for most computers. This particular technique represents the lowest cost solution.

Online UPS converts AC mains into DC before inverting again to AC to power the load with a synthetic sine wave. A battery connected across the DC link acts as the backup power source.

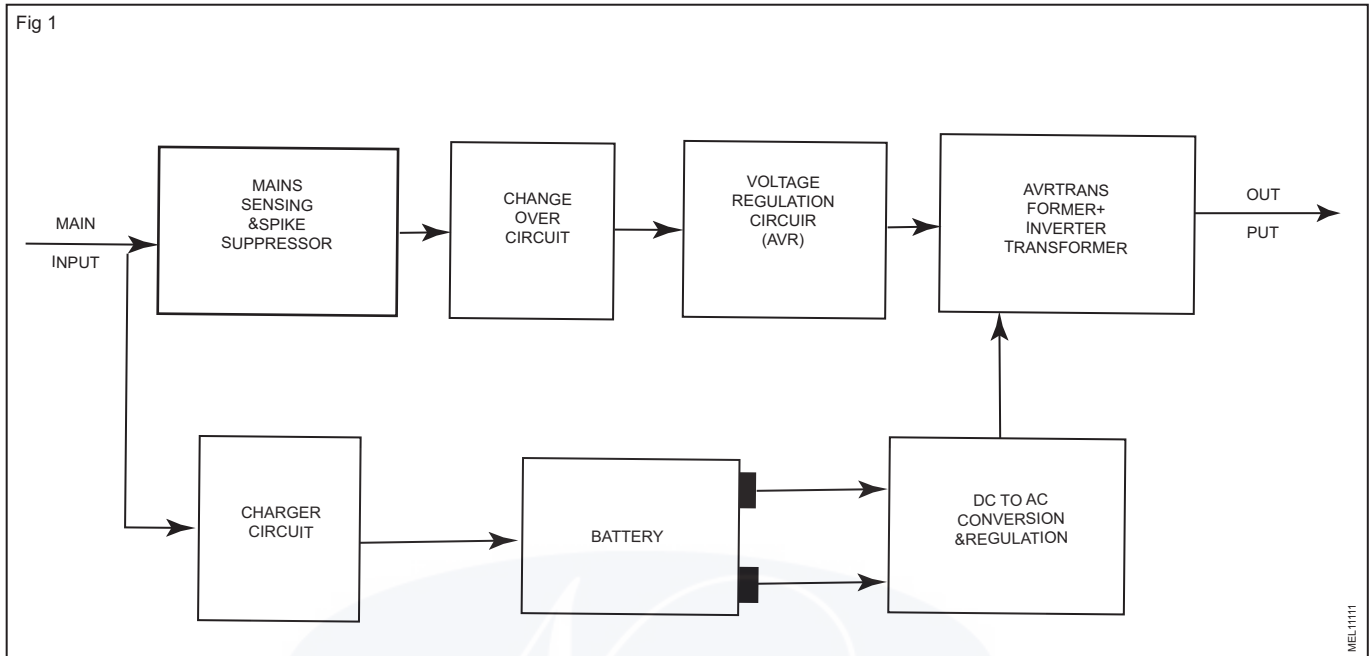
This gives a supply for the computer that totally isolates the input mains from the load, removing all mains noise and with no break when the mains fails.

Standby/OFF line block diagram

In the off line UPS, the load is connected directly to the mains when the mains supply is available. When outage/over voltage/under voltage conditions are detected on the mains, the off line UPS transfers the load to the inverter. When the line is present, the battery charger charges the battery and the inverter may either be shut down or will be idling. Thus an off line UPS transfers the load to the inverter. When the line is present, the battery charger charges the

battery and the inverter may either be shut down or will be idling. Thus in an off line UPS, there is a load transfer involved every time, the mains is interrupted and restored. This transfer is effected by changeover relays or static transfer switches. In any case there will be a brief period during which the load is not provided with voltage. If the load is a computer and the transfer time is more than 5ms, then there is a chance that the computer will reboot. Some modified designs incorporate a limited range of voltage regulation by transformer tapping and a certain degree of

transient protection by using RF filters and MOV's (Metal Oxide Varistor). Offline UPS is an economical and simple design and hence it is preferred for small rating, low cost units aimed at individual PC user's market. When the load is really a critical one an off line UPS is not acceptable. Usually square wave output off line UPS are available in market with lower loading capacities.



Advantage of OFF line UPS : high efficiency, small size, low cost.

Disadvantages : There can be change over complaint in offline UPS. Offline very much depends on battery. If battery fails entire system fails. Sometimes during changeover computer re-boots which causes loss of files. Another disadvantage is that output voltage will be a varying one. usually in the range of 200V-240V and hence not suitable to all gadgets.

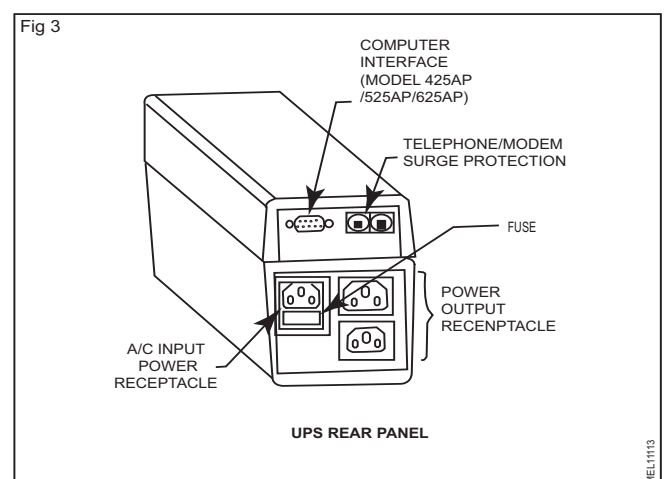
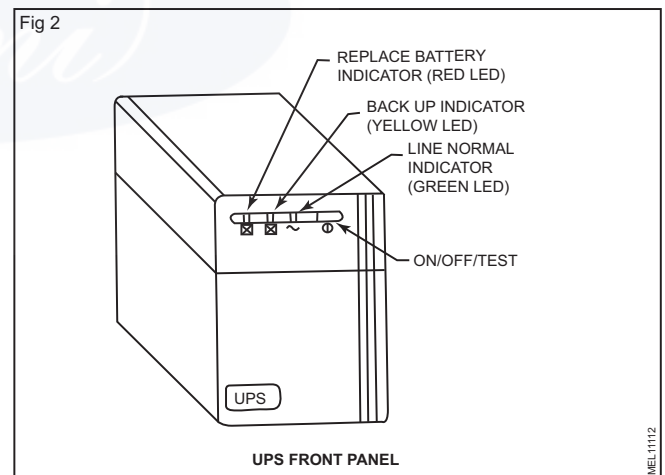
Different front panel indications and rear panel sockets/switches that are used in UPS

All UPS systems have

- Fuse/Fuse holder
- Switches
- Sockets
- Panel indicator (LED and Neon lamp)
- Meters (Volt/ampere)

Fig 2&3 shows the front and rear panel controls/sockets

Switches : On/Off switch and reset switch are commonly used in UPS. Reset switch is used to cut off an overload circuit and restart the supply. This is a push to off switch. In normal position this switch keeps the circuit on and when pushed, it cuts off the circuit.



Socket : A common 5 Amp. or 15 Amps . three pin power output socket is used in UPS to provide UPS output to the various devices. One can connect an ordinary 5/15 amp. plug to the UPS output.

Different LED indications/buzzers that are used in UPS

Main ON indication : It indicates mains input is present and UPS is working on mains.

Mains Low indication: It indicates mains input is low and is below a rated value.

Mains high indication : It indicates mains input is high.

Inverter ON indication : It indicates that UPS is working in the battery mode and mains is absent.

To get the output from UPS switch ON the 'Inverter ON' switch

UPS Trip indication : It indicates that UPS output is Off or tripped

Overload indication : Which indicates that the load current is above a pre-determined value.

Overload buzzer : It keeps whenever overload occurs.

Low battery warning : It indicates battery voltage is below a pre-determined value along the buzzer.

Battery charging indication : It indicates that battery is charging properly.

Output voltage low indication : It indicates that output voltage is below a pre-determined value.

General specifications & UPS protections

UPS are available from 500VA to 20 KVA or above. VA is volt ampere.

Power factor specification will be different for different manufactures. Suppose for 1 KVA UPS with a power factor 0.6 the load will be $1000 \times 0.6 = 600$ watts.

Normally a single PC takes around 180 watts. There are sine wave, square wave and quasi square wave output UPS. Usually sinewave output UPS is better than square wave output UPS.

General specifications

Output capacity = Output capacity will be in volt amperes (VA)

Input voltage = 230V AC + or - 20% , 50 Hz single phase sine wave

Output voltage = 230 V AC + or - 10% , 50 Hz square wave or sine wave

Battery = 7 AH, 12V Sealed Maintenance free(SMF) for OFF- Line (depends on the capacity of the UPS)

Tubular batteries from 40 AH to 160 AH(12V to 120V) for ON-Line (depends on the capacity of the UPS)

Availability of Automatic Voltage Regulation (AVR) feature.

Typical recharge time to charge 90% of the full capacity of the battery is 5 hours.

Different types of protection in UPS

Input fuse on mains : It protects the system from high voltage inputs, line disturbances and short circuiting etc.

MOV(Metal Oxide varistor) protection : MOV conducts when high input voltage appears thereby blowing the fuse.

Polyester capacitor for lightening protection: This is connected across the transformer winding. It burns when lightning occurs and protects the transformer.

Fuses to protect the MOSFETS : MOSFETS are highly sensitive to rapid changing currents. these fuses are used to protect the MOSFET.

Charger fuse to protect the charger circuit : If any fault in charger circuit occurs, fuse blows to protect SCRs.

Output high voltage protection MOV : This MOV is connected across output sockets phase and neutral. If feedback circuit fails the output voltage will jump to more than 300 volts. In such situation the MOV conducts to protect the load.

Overload protection : It protects the UPS especially MOSFET /IGBT when output current exceeds a preset value (overloading the UPS) . when this occurs, UPS output becomes OFF along with an indication.

Battery over charge/Discharge protection : It protects the battery from charging to a high value (SMF batteries will charge upto 15.8V) and tubular batteries upto 14.1V. It also protects the battery from getting discharged below a level(low battery protection). If the battery voltage is discharged below 10.5 V, then the UPS gets automatically switched OFF.

General tips for testing a UPS

Connect the battery to the terminals using a fuse wire. If any fault occurs in testing the fuse will blow to protect the UPS

Do the testings on no load condition

Check the gate voltages of the two MOSFET banks it should be the same. If PWM gate pulses are not present gate voltage will be around 5.6V . If the PWM gate pulses are present then the gate voltage will be around 2 - 2.5 volts

Some frequency meters are designed to measure pure AC frequency only. If the UPS output is square wave, then the reading will not be correct. To measure the correct frequency connect a 60/100W load at the output of the UPS. Then the frequency meter shows a near correct frequency.

For overload setting in On-line UPS, the load current is calculated by dividing the maximum load with the output voltage. This can also be measured using a clamp meter on the output terminal. Overload is set at this value of load current.

While using extension boxes either in the input or on the output of an UPS, ensure proper earth connection. Improper earthing may lead to poor line filtering and shock hazards.

If number of MOSFETs are connected in parallel, care should be taken to see that all the MOSFETs are of the same R_{ds} . For MOSFET R_{ds} value (drain to source resistance) and current rating are important.

Changeover in OFF-Line UPS system

The circuit diagram shows the output of OPTO coupler is given to a NAND gate IC, it controls the transistor BC548. This transistor controls the other transistor which controls the relay coil supply. Usually mains input is connected to the relay's common point. Output connection is taken from the normally open contact. When relay gets supply voltage across its coil, common comes in contact with NO contact. Normally closed contact is always left alone (There is no connection on this point). From normally open point the input to the ACR card is made.

In this type of relay control the battery voltage is used for relay coils. If the battery voltage is too low then relay coil will not get sufficient supply to trigger the switch. This may lead to the absence of mains voltage, even if mains is present and is healthy. This type of OFF-Line systems are battery dependent.

Some OFF-Line systems are battery independent. The coil supply is provided by the mains itself. Mains supply is reduced and rectified. This rectified supply is given to the changeover relay coil. Battery low voltage does not affect the relay coil supply. This type of OFF-Line UPS provides mains output irrespective of the battery condition.



Working principle of oscilloscopes

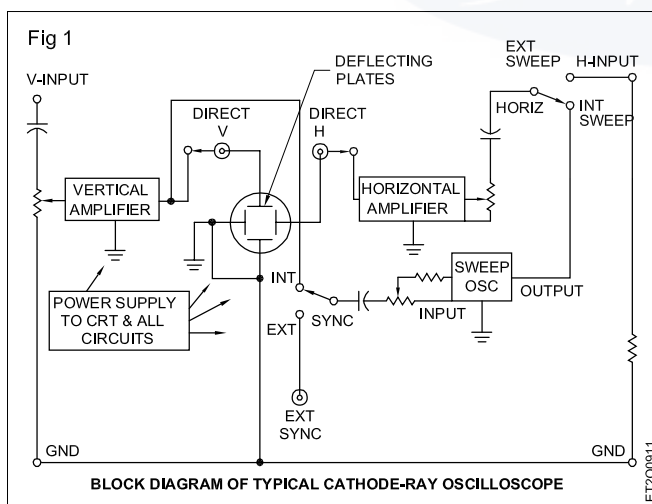
Objectives : At the end of this lesson you shall be able to

- list the five main functional blocks of a typical oscilloscope
- list the main parts of a CRT
- explain the need and method of post acceleration system
- explain the term Astigmatism
- explain the trigger and trigger method used in oscilloscopes
- explain how Lissajous figures can be used for frequency and Phase measurement
- explain the need of dual-beam oscilloscope and the main difference in the tube assembly used
- explain the need and principle of working of sampling oscilloscopes
- explain the working of storage tube in storage oscilloscope.

After using the Oscilloscopes in many of the previous exercises, you must be now familiar with the use of Oscilloscope. Hence, this lesson will focus on the working principle and internal details of a typical oscilloscopes used in laboratory tests.

Oscilloscope is an electrostatic measurement device. The movable pointer is composed of electrons being subjected to applied electrostatic field to be measured. As electrons are the lightest charged particles, they can directly respond to very fast changes of the electrical field strength. The heart of a Oscilloscope is the Cathode Ray Tube (CRT) which makes the signal visible by the action of an electronic beam.

Fig 1 illustrates the simplified functional block diagram of the internal parts of an oscilloscope.



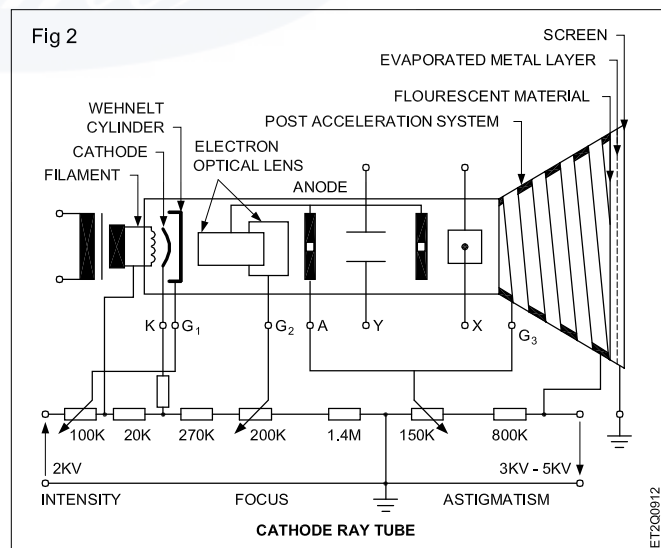
Referring to the block diagram of a general oscilloscope shown in Fig 1, observe that there are Five main functional units. These enable to display the input signal waveform.

- 1 Cathode Ray Tube (CRT).
- 2 The Vertical amplifier : This provide ample deflection of the input signal.
- 3 The Horizontal amplifier combined with the Sweep oscillator/time-base: This provide suitable time base deflection in synchronism with the input signal.

- 4 The Power supply: This is made-up of a low voltage section and a high voltage section. This supplies the required power to all the internal parts of the oscilloscope.

The Cathode Ray Tube (CRT)

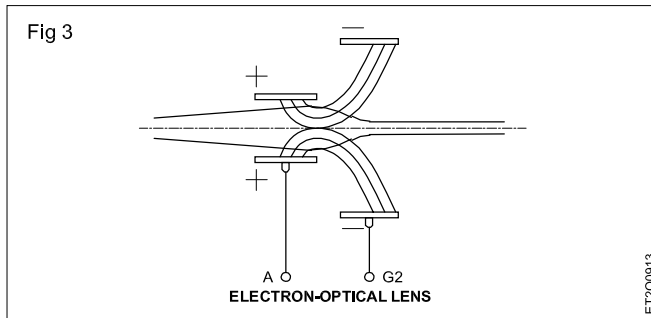
All the parts of the cathode ray tube is housed inside an evacuated glass enclosure as shown in Fig 2. CRT produces a narrow beam of electrons from an electron gun assembly,. This beam is then accelerated and impinges at high speed on the phosphorescent coating (screen) of the tube face. The impact of the narrow beam of electrons on the phosphor produces a single small spot of light on the screen. It is this small round spot of light that traces out the pattern of the wave form as it is moved by the vertical and horizontal deflecting plates of the CRT.



Cathode: The cathode is indirectly heated. Due to a barium oxide layer on the cathode surface, a temperature of around 800°C is needed to emit sufficient number of electrons per unit of time.

Wehnelt Cylinder: This is designed as an orifice plate. It is connected to a negative voltage towards the cathode which controls the beam and intensity with the help of P1. This way the brightness of the screen trace can be adjusted.

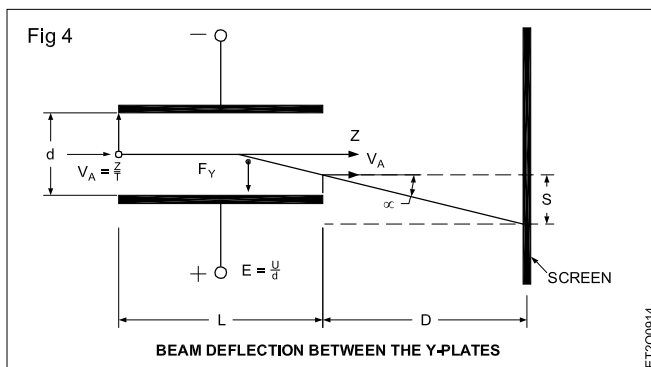
Electron Optical lens: This does the compression of the electron beam onto a thin line as shown in Fig 3.



The effect of the electron optical system quite similar to the effect of a convex lens on visible light. By changing the potential of G2 towards A with the help of P2, the focal length of the lens can be set to produce a depiction which is well in focus on the screen. After having passed the electron optical lens, a parallel beam is obtained having a circular cross-section of a very small diameter.

Anode: It is the last electrode of the beam generating system. It is supplied by a voltage which may range between 1 and 5KV with respect to cathode. This high potential accelerates the electrons to a considerable high speed. This high speed (kinetic energy) is converted into a visual trace on the screen. The center of the anode is designed as a hole, hence the electrons are channelled through it and continue to travel with a constant velocity before hitting the phosphor material behind the face of the screen of CRT.

Deflection Plates: After having passed the anode opening, the electrons approach the uniform fields of the deflection plates. First they reach the Y-plates which deflect the beam in vertical direction. After that they pass the X-plates designed quite the same way as the Y-plates but turned 90 degrees towards the Y-plate. This way they deflect the beam in horizontal direction. The voltage to be measured is applied to the Y-plates, which deflects the beam by an amount (say) S as shown in Fig 4. Further details of the CRT is discussed later.



Vertical Amplifier

The frequency response characteristics and sensitivity of the oscilloscope are determined chiefly by the vertical amplifier. Greater the sensitivity means narrower bandwidth. This is because, the product of gain times bandwidth is constant for a given amplifier.

Therefore the sensitivity of the oscilloscope is specified as so many volts/cm of vertical deflection at mid-band frequency. The 3db down response will also be specified at some specific band edge frequencies.

Sweep Oscillator

The sweep oscillator for generating the time base in a typical general purpose oscilloscope has a frequency range from say from 10Hz to 30 KHz. It is controlled by a coarse-frequency switch and a fine-frequency variable control. The scope can thus display as many as 5 cycles of a 50Hz input and 10 cycles of 300 KHz input.

Almost in all oscilloscopes the time base is synchronized, allowing a choice between "internal" and "external" synchronization. In the internal sync position, the fine-frequency control is adjusted to set the sweep frequency as close as possible to the desired sub multiple of the input signal, then a small amount of the synchronizing voltage from the input signal locks the pattern into a stationary display. Further details about triggering is discussed in subsequent paragraphs.

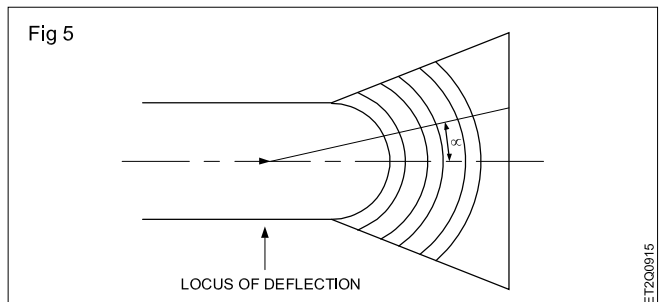
Horizontal Amplifier

The horizontal amplifier serves two purposes. In ordinary waveform presentation, it amplifies the time-base voltage from the sweep oscillator. A control makes it possible to adjust the width of the resulting pattern. When switched out of the internal sync position, it accepts an external signal applied to the horizontal terminals and amplifies the horizontal signal to the amount required for horizontal deflection. More details about the x-deflection is discussed in further paragraphs.

Post acceleration system

A high velocity of the electrons would certainly a cause bright appearance of the depicted spot on the screen. A high speed is obtained from a high anode voltage, which unfortunately would spoil a good deflection S. But both good deflection and a bright depiction need be coped up with.

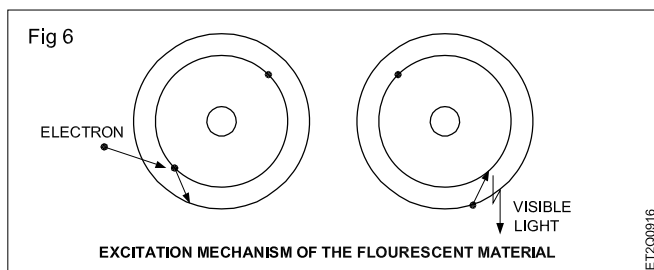
To achieve this a post acceleration system is introduced. This system does not effect the deflection angle α , but only provides additional acceleration to the electrons. For this purpose concentric rings made from graphite or a spiral are introduced into the conical part of the glass bulb as shown in Fig 5.



Their potential rises up to 5 kV. The rings provide areas of equal potential which are ball sections, having their centre at the locus of deflection. This design allows a post acceleration independently of the deflection angle. For any a, the electrons have to pass the same difference of potential resulting in the same final velocity for each of them. Eventually they hit the fluorescent screen with a very high speed thus producing a bright trace.

Fluorescent screen

The screen of an oscilloscope is covered from inside with a fluorescent substance which consists of zinc compounds. They differ from each other by different persistence and different colours. Once the electrons hit the fluorescent material they excite it for visible light radiation as shown in Fig 6.

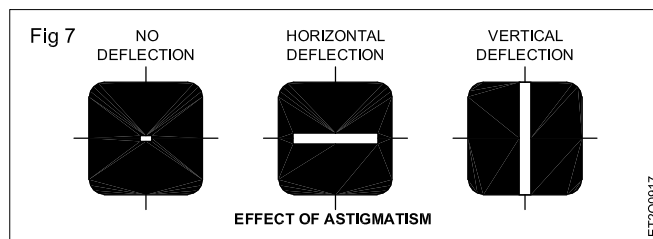


The high velocity electrons hit the atoms of the light emitting substance and they lift their electrons to orbits (around the atom core) which are of a higher energy level. They stay there for a certain short time which is typical for each substance. After that they fall back to their original orbit of lower energy level. The energy difference is emitted as a light quantum $h \cdot \nu$ ('h' being PLANCK's natural constant and ν , the frequency of the radiated light).

In order to avoid a negative charging of the screen trace (due to the charges of the arriving electrons) a thin metal layer of the thickness of only a few molecules is deposited on the screen from inside. It is transparent and grounded in order to draw out the incoming charge.

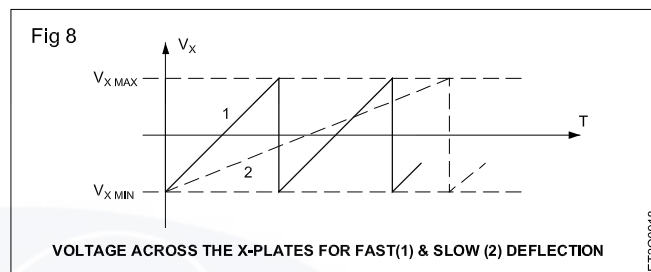
Astigmatism

The cross section area of the beam should be a narrow circle, and so should the light spot on the screen. For this purpose the electric fields of the beam generating components need to be extremely symmetrical and the deflection fields should be homogenous. Because of this condition the potentials right in the middle between the y- and the x- plates need to be the one of the anode as well as the one of the orifice plate between them which serves shielding purposes. The symmetry is set to its necessary amount with the help of potentiometer P_3 . In case P_3 is not properly adjusted the horizontally deflected beam may be depicted well in focus but the vertically deflected one may produce a wide screen appearance as shown in Fig 7. Modern scopes provide low tolerances for all components. So there is no need for any astigmatism potentiometer. But older types may still have it.



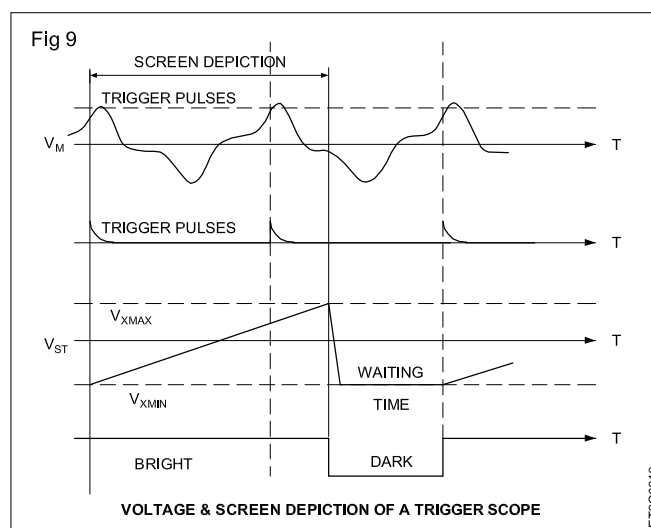
Time base

In order to deflect the beam horizontally proportional to time a voltage V_x is needed which rises linearly with time. It is used to control the position of the light spot on the screen in x-direction. Its lowest value $V_{x \min}$ should set the beam to the left edge of the screen and its highest one $V_{x \max}$ to the right edge. The whole run should be provided repeatedly. The switching back from the right to the left edge should be effected as fast as possible. So, a saw tooth X-deflection voltage as shown in Fig 8 serves this purpose. For trace 2 the deflection is slower than for trace 1.



Trigger method

The trigger instant of the saw tooth is linked to a certain instantaneous value of the measurand, the trigger level, and its slope. Once the trigger level is reached by V_m at a positive slope as shown in Fig 9, the beam starts at the left edge, for at that instant $V_{ST} = V_{x \min}$. The light spot travels with a certain speed in x-direction across the screen according to the slope of V_{ST} and finally reaches the right edge.



V_{ST} is equal to $V_{x\max}$ then. Immediately after $V_{x\max}$ was reached the beam is switched back to the left edge where the trace of the measurand during the positive slope of V_{ST} . No depiction appears during the fly back time and the waiting time of the spot.

The trigger “level” may be shifted throughout the whole voltage range between peak to peak value of the measured. Doing this the depiction appears to be shifted in x-direction.

The slope of the saw tooth voltage V_{ST} can be changed in wide range from very slow to very fast allowing to depict many periods of the measurand on the screen but also only part of it.

Power supply

Different voltages required for the oscilloscope is obtained from this unit. The exact value of voltages required for each model of the oscilloscope depends on the type of CRT used, the type of circuit components used and so..on. However essentially any CRO need a high voltage supply for the CRT and low voltage supply for the electronic circuits.

Lissajous figures

Lissajous figures are depictions of two voltages applied to the Y- and X- plates which are periodical time functions. But the depiction on the screen is independent of the time. Whether the frequencies of V_x and V_y change fast or slow is not decisive. The trace will be the same in any case.

Frequency measurements

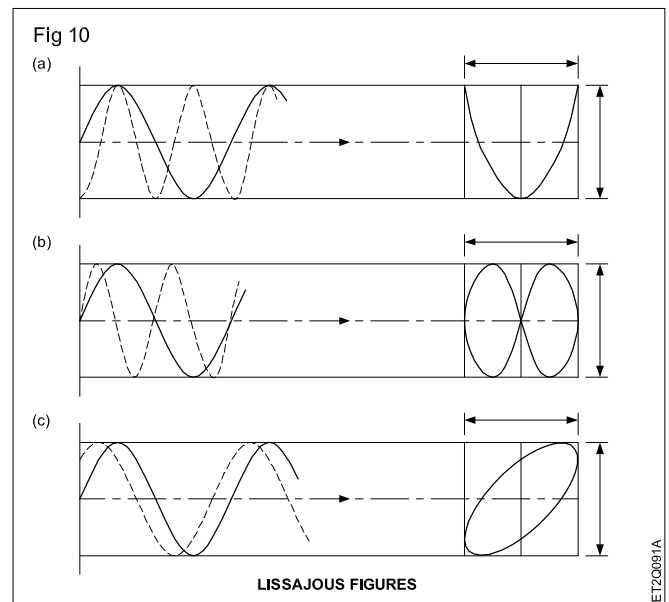
Lissajous figures can be obtained from independent voltages V_x and V_y . In this case they are usually used for frequency measurements. The unknown frequency can be determined if the other one is known. Consider Fig 10a, the first graph depicts a voltage V_x of frequency f_x and a voltage V_y of the frequency $f_y = 2f_x$. The Lissajous figure can be sketched with the help of the voltage pairs V_x , V_y for the 9 sample points.

The same result for the frequency f_y is obtained from the second and third illustrations at Fig 10b and Fig 10c. But here the Lissajous figure appears quite different due to the phase shifted U_y . By these figures, if one frequency is known the other can be determined. But the ratio of f_y/f_x needs to be a proper fraction in order to obtain a standing screen depiction. This condition can be met with the help of a tunable frequency generator for f_x . As long as the frequency condition is not fulfilled the depiction changes its shape continuously.

Frequency measurements employing the Lissajous method have lost importance because electronic frequency counters provide easy handling possibilities and high accuracy.

Phase measurements

If the Lissajous figure is one continuous loop as in the third depiction of Fig 10c. The frequencies f_x and f_y are equal. In this case the figure can also be used to determine the phase angle j between V_x and V_y .



ADVANCED OSCILLOSCOPES

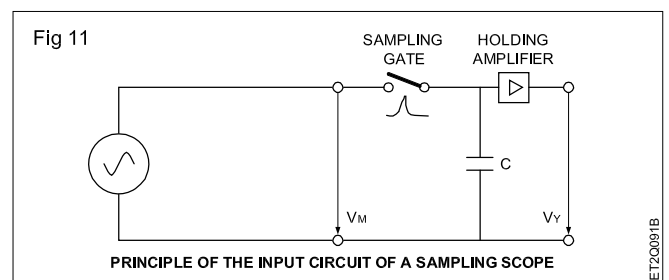
Dual Beam Oscilloscope

In circumstance while measuring high frequency signal a dual trace may not be fast enough to display in either its “alternate mode” or “chopping mode”. For such application the oscilloscope need to have two separate electron guns in one tube assembly. Such oscilloscopes are more expensive than general single beams oscilloscopes.

Sampling oscilloscopes

To extend the range of oscilloscopes to hundreds of megahertz, a highly effective method is used called the sampling method. This method looks at the successive points with each sweep. In this way the sweep need recur only in kilohertz range to investigate and display signals in mega and gigahertz range. By this method even though the display produced may be dotted in form, it is possible to visualize waveform patterns reasonably well.

Measurands of very high frequencies (> 100 MHz) cannot be directly depicted using the trigger method. But the sampling method can cope up with such signals in case they are periodical. The measured voltage V_m is continuously scanned by taking a sample, during each cycle. But the time instant of measurement is slightly shifted consecutively for each new sample. The principle of the input circuit of a sampling scope is shown in Fig 11.



The sampling gate is periodically activated for a very short sample time, scanning the voltage V_m which is provided by its source generator. During the sample time the capacitor C is charged. It takes over a voltage equal to the instantaneous value of V_m during the sample time. After the sampling gate has opened, the charge of the capacitor remains unchanged and so does its voltage. It is held constant until the gate is switched-on again and another V_m may change the capacitor voltage. This means: during the pause between two sample pulses, the capacitor voltage stays constant and an amplifier provides it at a higher (energy) level as V_y .

The rising signal V_m to be measured causes a trigger pulse periodically as shown in Fig 11. This in turn generate sampling pulses with a delay as shown in Fig 12. Each sample pulse connects the capacitor C for a very short time to the measurand and takes over its instantaneous voltage. Hence V_y changes in steps following V_m . If there are 10 sample pulses during 10 period durations of V_m , the voltage V_y has followed in steps only once. The x-plates are supplied by a step voltage arising in equal steps from $V_{x\min}$ to $V_{x\max}$. In this way the beam is deflected in steps from the left edge to the right edge. The step width is $1/10$ of $V_{x\max} - V_{x\min}$. In this way the light emitting spot on the screen leaps from one depicted dot to the other. A dotted trace occurs. It gives quite a good imagination of the actual signal.

The number of sampling pulses needed for a full depiction of one cycle, times the period duration of the measurand, equals the time for depicting the trace once. The screen shows a time expanded appearance of V_m . It is a stroboscopic presentation of the measurand.

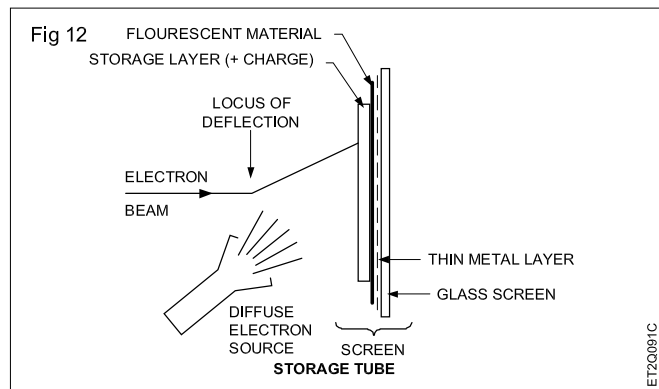
Storage Oscilloscope

The ability to store a display on the oscilloscope screen has great utility for observing the transients and other phenomenon in cases where it is not practicable to relay on photographic the trace at critical moments or in a rapidly changing circumstances. The storage oscilloscopes makes it possible to store repetitive signals from very low rates up to several hundred kilohertz and also single transients. Images of high contrast can be stored for several months. Storage oscilloscopes can provide readout for computers, printers and plotters such that leisurely examination of any waveform is possible.

CRT of Storage Oscilloscopes

Storage tube: Light emitting fluorescent substances are available with wide ranges of their persistence. But fluorescent materials with a very long persistence cannot cope with non-repetitive measurement voltages, which are available only once. The screen depiction may emerge too fast for proper assessment. In such cases storage tubes are of good advantage. The principle of action is shown in Fig 12.

Behind the screen there is additionally a storage layer, which is positively charged. Once the incoming electrons reach this layer, they neutralize it at the location of the trace to be depicted (due to their own negative charge).



According to the function $y(t)$ which the electron ray marks, a neutral trace is left behind amidst the positively charged area of the storage layer. This happens within the 'writing' mode of the scope which stores the information this way. For watching the trace a second source needs to be employed which emits electrons in a diffused way with comparatively low kinetic energy. These electrons (uniformly distributed) are sprayed at the storage layer. At the location of the neutral trace they may pass and can excite the fluorescent layer for light emission. The trace becomes visible. At the other positively charged area the electrons cannot pass. Their negative charge recombines with the positive charges and this part of the field remains in dark. At first a bright trace (or spot) appears within a dark surrounding producing a beautiful contrast. But gradually, more incoming electrons discharge the positive area of the surrounding field and finally they all manage to pass over to the screen thus exciting light emission everywhere. The contrast emerges and the stored information disappears gradually, because the whole area gets neutralized eventually. This happens within the scope mode of 'view'. Usually a short time of viewing may be sufficient to assess the measurand. So one should switch over to the mode 'store' in order to sustain the information. The storage time ranges up to 60 hours for storage scopes these days, even for the switched off device. Advances in the technology of insulation materials provide this feature. In case a new trace should be stored, the whole layer needs be 'erased' first. This is effected by recharging its entire area positively.

Storage tubes are expensive. The most recent oscilloscopes store the information in digital memories. The most recent of storage oscilloscopes are the *digital storage oscilloscopes*. Refer books listed at the end for information on digital storage oscilloscopes.

General Accessories of Oscilloscopes

Marker Generator

The marker generator provides visual time-interval or frequency reference marks on the scope trace. Such markers can select time intervals from 1 microsecond all the way upto 5 seconds. These are particularly useful in rise-time and pulse-duration measurements.

Electronic switch

The electronic switch displays two signals simultaneously on single-trace scopes. It is essentially a square-wave generator. The two signals which are viewed are each connected to a separate grid of a dual-section amplifier tube, operating with cutoff bias. The square wave is applied in such a fashion that is alternately unblanks each tube section, thus presenting each signal alternately to the vertical amplifier of the scope, with appropriate vertical spacing between the signals. When the switching rate is high enough, each signal appears continuous, even though it is actually being periodically interrupted. Switching rates as high as 100 kilohertz are available.

Oscilloscope probes

The probe connects the test circuit to the testing instrument, without appreciably altering, loading, or otherwise disturbing the test circuit. Although probes may be given many different names, they fall into three principle types:

- 1 Direct probe (or test cable)
- 2 Circuit-isolation (or voltage-divider) probes
- 3 Detector (or demodulator) probe.

Direct probe

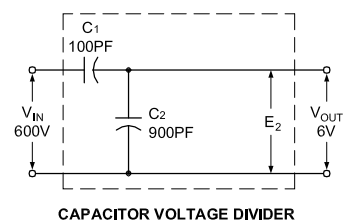
This simplest probe is a shielded coaxial cable. It avoids the stray pickup that can be troublesome when low-level signals are being examined with correspondingly high-gain settings on the scope. It is most effective in relatively low-frequency circuits, where cable capacitance is not too important, and in such cases allows the maximum sensitivity of the scope to be used. However, the shunt capacitance of the probe and cable is added to the input impedance and capacitance of the scope, and lowers the scope's response to high impedance and high-frequency circuits being tested. Where this effect becomes important, the isolation probe is used.

Isolation probe

To avoid the undesired circuit loading, the isolation probe is used to decrease the input capacitance and increase the input resistance of the oscilloscope. The low-capacitance probe attenuates the scope's input capacitance usually by a 10:1 ratio, necessarily attenuating the input signal by the same ratio. To equalize the time constant, resistors are commonly placed across the voltage-dividing capacitors.

Where it is desired to measure signals of high enough voltage to damage the scope's input circuitry, signal attenuation is desired, and a high-voltage probe is used. Such a probe consists of a resistive or capacitive voltage-divider circuit as shown in Fig 13. These probes allows only a small portion of the input signal to reach the scope, but preserves the waveform accurately. When isolation probes are used to measure the magnitude of test voltages, the attenuation factor by which the probe divides the input signal voltage must obviously be known, and measurements calculated accordingly.

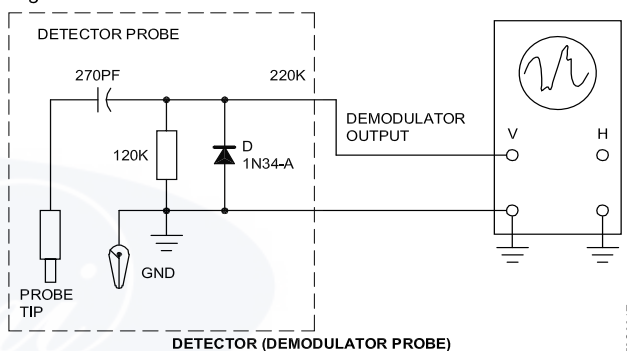
Fig 13



Detector or Demodulator probe

The detector or demodulator probe is used in analyzing response to high-frequency modulated signals used in communication, as in the R-F section of AM, FM and TV receivers. By rectifying and by-passing the radio frequency, the modulation envelope is displayed on the scope. This allows a scope capable of only audio-frequency response to trace communication signals in the range of hundreds of megahertz a range which would otherwise be far beyond the capabilities of all but highly specialised scopes. The detector probe separates the audio-frequency modulation component from the radio-frequency carrier.

Fig 14



The detector circuit of the probe is shown in Fig 14.

After being rectified by diode D, the audio signal corresponding to the original modulation is passed on to the output terminals. The resulting display of the audio signal can then be used to indicate the proper alignment and response of each tuned R-F circuit.

Summary of major applications of oscilloscopes

It would be impractical even to attempt a complete listing of all the applications of the oscilloscope. However a few major fields of application are listed below:

- 1 Measurements of instantaneous voltage and of current (indirectly)
- 2 Amplifier gain, frequency response and phase relations
- 3 Rise time and duration of pulses
- 4 Modulation percentage
- 5 Distortion
- 6 Timing comparisons of pulse trains as in ignition systems, vibration studies, telemetry and digital techniques.
- 7 Response of tuned circuits
- 8 Industrial display of non-electrical quantities, using transducers for mechanical, thermal, optical, acoustic, chemical, biological and nuclear properties.

Measurement using dual trace oscilloscope - 1

Objectives : At the end of this lesson you shall be able to

- procedure to calibrate the given CRO using internal calibration signals
- procedure to measure frequency and phase of a signal using Lissajous waveform method
- procedure to measure the propagation delay in digital signals using a divide-eight-counter.

This class room session is expected to be highly interactive and brainstorming. In this session, the instructor should take-up each of the objective listed above separately and guide the trainees to develop a procedure for carrying out the task. For example, in this classroom session, the instructor should first take-up the first objective “**procedure to calibrate the given CRO using internal calibration signals**” and brief the trainees the nature of task (calibration of CRO).

The instructor should then divide the class into 4 groups and instruct them to draft the procedure to carry out the task in hand (“to calibrate the given CRO using internal calibration signals”). To aid the trainees work, they should be provided with copies of the oscilloscope manuals, related reference books (available in the library) and advised to refer previous lessons on oscilloscope. With these reference materials in hand and the demonstration witnessed by them in the previous exercises, the trainee groups should draft the procedure for carrying out the task in hand (each group should develop one draft).

The draft developed by each group should be discussed with the entire class. During the discussion, the trainees should be motivated to point out procedural errors in the drafts and suitable correction to it. After discussing all the drafts (4 drafts in a class of 16 trainees), the instructor should generate a procedure taking all vital points from the drafts. This shall be used as the final procedure for carrying out the task in the laboratory.

The same procedure should be repeated for all the tasks given under the objectives for this lesson.

NOTE TO TRAINEES

- Do not waste time in drawing beautiful/colourful sketches of the test setup. Just rough pencil sketches will do.
- Refer to books and or consult your instructor when you are in doubt.
- Give opportunity for every trainee in your batch to give his idea.
- Jot down ideas given by the group members. This will help you finalize the draft more easily and in less time.

Instructor to refer IG for more tips on conducting this brainstorming session.

Controls and functions in oscilloscopes

Objectives : At the end of this lesson you shall be able to

- explain the use of X5 and X10 magnification of in the vertical deflection factor control
- explain the use of different coupling modes of the vertical input.
- explain the use of CH1 and CH2 inputs of a dual trace oscilloscope effectively
- explain the use of Alternate and Chopped modes for two inputs
- list the use of different triggering modes such as INT, EXT, EXT/5 or EXT/10
- list the use of different trigger source coupling modes such as LF reject, HF reject
- list the use of different sweep display modes
- explain the use of X-Y mode of operation
- explain the use of Z -axis input.

INTRODUCTION

In addition to the standard front panel controls of a general purpose oscilloscopes, certain of the controls and functions which are essential while displaying the measurand clearly are discussed in this lesson. Also some tips while using the oscilloscope are also discussed in this lesson.

Focus and Intensity

When the oscilloscope is switched on with the power on switch, the first thing to do is to get a beam trace on the oscilloscope screen. Focus and intensity controls together help to get a sharp, low intensity trace. Lower intensity not only allows to focus the display to a very fine trace, but also increases the life of the CRT of oscilloscope. The trace intensity should never be so bright that it burns a hole in the phosphor coating on the CRT screen.

The damage to the CRT with an extra bright trace is much more severe, particularly when you are working at slower sweep speeds.

Astigmatism

Some oscilloscopes have astigmatism control that should be so adjusted that the focus control is effective on the horizontal and vertical portions of the trace. Simultaneously, astigmatism control should be adjusted with a pulsed waveform displayed on the screen.

Trace rotation

It can be used to make the beam trace perfectly horizontal in the absence of any input signal. It is usually a trimmer whose adjustment screw can be seen on the scope's front panel or on the rear panel.

Beam find

Often we come across a situation where we have switched the oscilloscope on, increased the intensity level, selected the auto sweep mode and tried to adjust the horizontal and vertical position controls but still have not been able to see the beam trace. Beam find control can be used to locate the beam irrespective of where it is. Pressing this button compresses the range of horizontal and vertical position controls and the result is a dot somewhere on the screen.

Keeping the button pressed, adjust the two position controls to bring the dot to the centre of your scope's screen. Release the button and you will see a trace right in the middle of the screen.

Horizontal and vertical position

Horizontal position (indicated on some scopes as <---> and vertical position (indicated on some scopes as) are used to shift the trace horizontally and vertically respectively. There is usually a common horizontal position control in a dual trace oscilloscope. The position control shifts both the traces in the horizontal direction simultaneously. However, there are two separate vertical position controls for the two channels.

Calibration

All oscilloscopes have a CAL output. The amplitude and the frequency of the calibration signals are indicated on the front panel by the side of the output. The calibration signal can be used to check the amplitude and the time base calibration of the oscilloscope.

Some oscilloscopes provide two calibration signals, both having the same frequency but different amplitudes. Oscilloscope may have two calibration signal outputs i.e. 2Vp-p at 1 kHz and 200m Vp-p should be checked with both the signals. Scope's calibration should be adjusted at regular intervals.

In some oscilloscopes, the output of calibration is indicated by a glowing LED. You will find an LED near the time base setting and LEDs near the vertical deflection factor selector switches of the vertical input channels. Calibration signal is also employed to adjust the probe. The conditions of an under compensated or an overcompensated probe can be easily seen with the calibration signal used as a reference.

Bandwidth limit

Many high sensitivity, high bandwidth oscilloscopes have bandwidth limit control. Though higher bandwidth capability lets you capture high frequency signals, the unwanted high frequency noise also creeps in. It is particularly troublesome when we are viewing a very low level signal (say a few

millivolts) of moderate frequency. Due to high bandwidth capability of the scope, the desired signal is often seen accompanied by a lot of hash.

Volts/div and time/div controls

Volts/div and time/div are the controls that need frequent adjustment while viewing and analysing signals. While the former selects vertical sensitivity and is set as per the amplitude of the signal to be viewed, the latter sets sweep speed and its setting is governed by the signal frequency. Both these controls have a selector switch setting and a fine control. The fine adjustment control in both cases should be kept in the calibrated position. The selectable positions in case of these controls are in the decades of 1-2-5.

In most oscilloscopes, there is provision for X5 magnification in the vertical deflection factor control which makes the oscilloscope more sensitive by a factor of 5. That is, 5 mV/div to 5V/div range becomes 1mV/div to 1V/div. But then we must always remember that this enhancement in vertical sensitivity is at the cost of reduced accuracy. Accuracy specification of typically ± 3 percent may deteriorate to ± 5 percent. This magnification is usually obtained by pulling the fine adjust control knob in the vertical deflection factor selector switch.

Similarly, a magnification of X10 is usually available in the time base setting, which means that sweep speed at any setting can be increased by a factor of 10 by using this feature. This enhancement is also at the expense of degradations in sweep speed accuracy. The change in accuracy may again be from ± 3 percent to ± 5 percent. X10 magnification is also achieved by pulling the fine control adjust knob in the base selector switch.

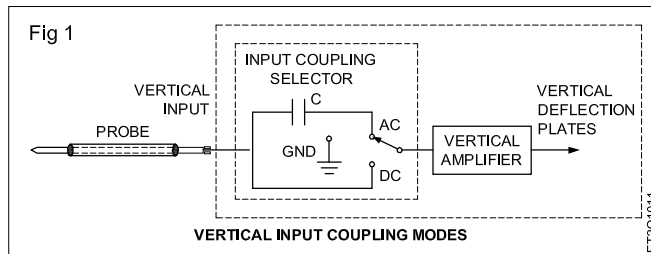
In some oscilloscopes, the time base selector has two switches and a fine adjust. One of the two switches, selectable by bigger of the two knobs, is used to select the main sweep speed. There is another switch concentrically located with a smaller knob. This is used to choose the delayed sweep speed. This second rotary switch is present only in oscilloscopes having delayed sweep facility. Also, the two switches are so internally arranged that the delayed sweep speed can never be set to be slower than the main sweep speed.

Input coupling

The coupling selector is a three-way switch, to select either DC or AC coupling and ground. In DC coupling, the input signal is fed directly into the amplifier, while AC coupling enables blocking of the DC component of the input signal and passes only the AC component of the signal to the Y amplifier. In the ground position, the input of the Y amplifier is grounded. Hence, care should be taken to ensure that the input signal is not grounded in the ground position and that only the input point of the Y amplifier in the oscilloscope is grounded.

Vertical input coupling modes

All oscilloscopes have two vertical input coupling modes, namely AC coupling and DC coupling as shown in Fig 1. In DC coupling selection, the signal to vertical input BNC receptacle is directly routed to the input of the relevant vertical amplifier as shown in Fig 1 inside the scope. As a result, what you see on the oscilloscope is what you feed into it.



The DC coupling mode is used in majority of oscilloscope measurements whether it is measuring DC amplitudes or seeing logic low and high levels over analysing transient and repetitive AC waveforms over the specified bandwidth of the oscilloscope. However, when it comes to measuring only the amplitude of a certain DC voltage with no intention of analysing the quality of DC or looking for presence of any noise spikes, the oscilloscope in the DC coupling mode does the job.

In the AC coupling mode, the applied signal is routed to the vertical amplifier input through capacitor (Fig 1) with the result that DC, if any, in the signal gets blocked and only the AC or the time varying part is able to get through and reach the vertical amplifier input. So the displayed waveform is not what you actually feed. For instance, if you want to analyse noise spikes or ripple content riding on a DC you would have no option other than going in for the AC coupling mode.

In the DC coupling mode, the beam would go off the screen as you increase the vertical sensitivity to get an expanded display of comparatively much smaller ripple amplitude. In the AC coupling mode, you could expand the display and make the ripple portion fill the entire screen for detailed analysis.

There is a ground position (designated GND) available on the coupling selector. In this position, the input of the vertical amplifier is grounded and this position can be used to know the position of the beam for zero input.

Input impedance

This is the impedance at the Y input point and is normally specified as 1 M ohms shunted by 25 pF. It is actually the effective resistance and capacitance across the Y input. All oscilloscopes have a standard input impedance of 1M ohm paralleled approximately by 25 pF.

Maximum input voltage

It is the maximum voltage that can be safely applied to the Y input of the oscilloscope. For example, a model specifying the maximum input voltage to be 400V (DC + peak AC) means that the voltage of the signal to the input cannot exceed 400V, which includes both the DC voltage and the peak AC voltage of the signal.

Vertical operating modes and relevant controls

In a dual trace oscilloscope, if the two vertical input channels are designated CH1 and CH2, the available vertical operating modes usually are CH1, CH2, ALT (alternate), CHOP (chopped) and CH1 + CH2. CH1 mode selection implies that the beam traces the waveform applied to the channel-1 vertical input every time it sweeps across the screen.

With CH1 + CH2 mode selected, each sweep across the screen traces channel 2 vertical input waveform. When CH1 + CH2 mode (also referred to as ADD mode) is selected, what we see on the screen is sum of CH1 and CH2 signals as a function of time. Alternate (ALT) or chopped (CHOP) modes are selected when we intend to see two different signals simultaneously.

Alternate or chopped

ALT and CHOP modes are used in two different situations. In the ALT mode, CH1 and CH2 signals are traced on alternate sweeps, i.e. if n th sweep traces CH1 signal then $(n+1)$ th sweep would trace the CH2 signal, $(n+2)$ th would trace the CH1 signal again and the process would continue. If the sweep speed is low, say slower than 10 ms/div or so, we will see a blinking display of the two sweeps. For faster sweep speeds, the two displays appear to be present at the same time. The ALT mode display of two channels thus gives an uncomfortable display when the signal frequencies are low. This mode should preferably be used for viewing high frequency signals.

In the CHOP mode, each sweep across the screen switches the beam between CH1 and CH2 at a very fast rate (the chopping frequency is typically 50 kHz to 100 kHz). In fact, we can see this chopping effect by selecting the CHOP mode and choosing a time base setting faster than the chopping frequency. CHOP mode is not suitable for viewing very high frequency signals as you are likely to miss vital signal information during the time period when the sweep is tracing the other signal. CHOP mode is, however, the right mode to select for viewing signals having frequencies of a few kilohertz or more.

In some oscilloscopes (usually the ones with lower bandwidth) we do not have a separate select button for CHOP and ALT modes. Instead, we have the dual mode in which the oscilloscope has in built circuitry to give a chopped sweep operation for lower frequency signals (or slower time base settings), and an ALT mode for viewing high frequency signals (or faster time base settings). The range of time base setting for which the scope offers a CHOP mode or an ALT mode is usually indicated on the time base selector switch.

In the front panel of a oscilloscope you would notice a light coloured semi-circular band from 0.5 s/div setting to 1 ms/div setting indicating the CHOP mode and another dark semi-circular band from 1 ms/div indicating ALT mode.

LF Rej

This is a method of coupling the trigger signal with the trigger circuit. The trigger signal is fed to the trigger circuit via a high-pass filter, where the low frequency component (less than 10 kHz) is eliminated. Thus, triggering is effected only by the high frequency component. When the trigger signal contains low frequency noise (particular hum) it is eliminated so that the triggering is stabilised.

HF Rej

In this method, the trigger signal is fed via a low-pass filter where the high frequency component (more than 30 kHz) is eliminated. Triggering is effected only by the low frequency component.

Triggering modes and relevant controls

All modern oscilloscopes are triggered sweep oscilloscopes, i.e. each sweep across the screen is initiated by a trigger signal either generated inside the scope or supplied externally. The source of trigger signal, the way it is coupled and the controls like 'trigger slope', 'trigger level' and 'trigger hold off' enable you to make full use of the equipment and get a stable display of many a complex waveforms or trigger on the most elusive transient events.

Source of trigger signal

This first relevant control is the one that selects the source of trigger signal. The available options in most of the oscilloscope are internal (INT) line, external (EXT) (Refer to the top extreme right and bottom extreme right of the oscilloscope's front panel in Fig 1).

When we have selected the INT source of trigger, the trigger signal is generated from the signal to be viewed. A small part of the vertical input signal is taken off, amplified, shaped and then treated as the trigger signal. In a dual channel oscilloscope, where we have two vertical inputs, a separate control decides whether it is a part of CH1 signal or CH2 signal that is to be used for generating the trigger signal. Here, if we select ALT, the trigger signal source is according to the vertical mode displayed. We should also remember that selection of CH1 signal or CH2 signal or ALT trigger arises only when trigger source selection is on INT.

When the trigger source is line, the oscilloscope picks up 50 Hz signal from its power transformer and uses this for producing trigger signal. It is suitable for getting a stable display of signals having power line frequency like ripple on a power supply.

In the EXT mode, the trigger signal is applied externally. The trigger signal amplitude requirements are specified by the manufacturer. Some scopes also have EXT/5 or EXT/10 trigger inputs. The trigger signal applied to this input is alternated by the given factor before it is applied to the trigger circuit. This mode is used when the external trigger signal level is too high.

Trigger source coupling mode

The coupling mode selector determines the way the trigger signal is coupled to the trigger amplifier. The available options on most of the 100 MHz oscilloscopes are DC, AC, Low Freq Rej (low frequency reject), High Freq Rej (high frequency reject) and TV. The Low Freq Rej coupling mode is usually not present in lower band-width oscilloscopes (upto 50 MHz bandwidth).

In DC coupling of trigger source, the trigger signal is directly coupled to the trigger circuitry. This mode is used when triggering is required to be effected including the DC component of the trigger signal. It is suitable for viewing DC and low frequency signals.

In AC coupling, the trigger signal is AC coupled to the trigger circuit. This is the most commonly used trigger source coupling mode as stable triggering can be achieved without being affected by the DC component of the input signal.

In the Low Freq Rej mode any frequency component below a few kilo-hertz present in the trigger signal attenuated. This mode should be used when low frequency components, 50 Hz hum for instance, is present in the trigger signal. High Freq Rej mode is used when any high frequency components present in the triggering signal are creating problems in getting a stable display. In this mode, high frequency components greater than 50 kHz present in the trigger signal are attenuated.

The TV coupling mode is used exclusively for viewing TV video signals. The signal is AC coupled to the TV sync separator circuit. The sync separator picks up the sync signal which is then used as the trigger signal. With this mode we can obtain a stable display of TV video signals.

Trigger slope and level

Trigger slope selection determines the slope of the trigger signal that triggers the sweep. When we select a (+) slope, the sweep is triggered anywhere on positive going or low-to-high transition of the signal. In case of (-) slope, the sweep is triggered anywhere on the negative going or high to low transition of the signal.

The trigger level decides the signal level (positive or negative) where the triggering takes place. If the signal has both positive as well as negative amplitudes, we can trigger on a positive slope and a negative level or a negative slope and a positive level as well. When we select a positive slope, the waveform can be triggered anywhere on the positive slope of this waveform, i.e. from negative peak towards positive peak. The level can be either negative or positive. Similarly, when we select a (-) slope, the waveform can be triggered anywhere on the negative slope, i.e. from positive peak towards negative peak. The level can be either be positive or negative.

Trigger hold-off control

This control can be used to adjust the pause between initiation of two successive sweeps and is particularly useful for viewing signals that do not repeat symmetrically. In the absence of trigger hold-off feature, it may be difficult to get a stable display of waveform of this kind. The trigger hold-off control can be used to trigger the sweep at the right time.

Sweep modes and relevant controls

The first selection that we have got to do is that of the sweep triggering modes. Usually, three modes are available on almost all oscilloscopes. They are auto (automatic), normal and single sweep modes.

In the auto sweep mode, the sweep generator is a free-running oscillator if there is no triggering signal, internal or external. That is, if the trigger source has been chosen to be INT, we will see a beam trace even in the absence of any vertical input. When a triggering signal is applied, the scope becomes a triggered sweep one and the trigger signal initiates the sweep as per slope and level settings. The auto mode is quite convenient when we are interested in seeing DC voltages or simple waveforms.

In the normal sweep mode, the triggering signal only initiates the sweep. In the absence of any trigger, we do not see any trace on the oscilloscope screen. In the normal mode, we have to carefully select the slope and adjust the level to get a display of the signal. This mode is suitable for viewing complex waveforms and single shot events.

In the single sweep mode, when a triggering signal is applied, the first genuine trigger initiates a sweep and after that all subsequent triggers are ignored. So there is only a single sweep. When the single sweep mode is selected, the oscilloscope gets ready to receive the trigger. This mode is very useful for viewing single-shot events.

Sweep display modes

The second selection that needs to be done is that of the sweep display mode. The available choices are the main sweep, delayed sweep, intensified sweep, triggered delayed sweep. These may be designated as A-sweep (mainsweep), B-delayed sweep (delayed sweep), A-intensified (intensified sweep) where the two input channels are referred to as A and B.

The main sweep is what we have been referring to so far. Its speed is set by the main time/div selector switch. It is suitable for most measurements. But what happens when we want to view a small part of a comparatively lower frequency signal on an expanded scale to look for noise glitches? If we try to expand the time base, the desired portion on the waveform is likely to go off the screen and all our efforts to bring it to the centre of the oscilloscope screen with the horizontal position control are rendered useless. One method to overcome this is to use X10 magnifier available with the main sweep. Engaging the magnifier expands the time base by a factor of 10 around the centre of the screen with the result that the desired portion stays on screen. This process is known as magnified sweep.

Magnified sweep too has its own problems. First, the intensity of the sweep diminishes quite a bit on expansion and second, this expansion may not be sufficient to permit a view of very fast glitches, for instance, a few nanoseconds wide glitch sitting somewhere on a waveform with a time period of a few milliseconds.

Delayed sweep is what comes to our rescue in such cases. As mentioned earlier, we have at our disposal two independent time base settings, one for the main sweep and the other for the delayed sweep. To make use of the delayed sweep facility, set the delayed time base at a much faster speed than the main time base. There is also a delay time multiplier (a multiturn potentiometer) control on the panel. Set that to the centre of its range. Engage the intensified sweep button. We would notice a small portion of the waveform being viewed on main sweep getting intensified. This implies that we have engaged the delayed sweep. The width of this intensified portion depends upon the time base setting of the delayed sweep.

The photograph is for a delayed sweep of 5ms/div. The width becomes narrower as we make the sweep faster. Thus, faster the delayed sweep, narrower is the intensified portion and larger is the magnification that we get. The position of this intensified portion is as per the part of the waveform we wish to expand.

After having adjusted the two things, engage the delayed sweep mode. The intensified portion fills the entire screen. In this mode, we can achieve much higher magnification without sacrificing the intensity. In some scopes, there is a provision for viewing the main sweep signal and the intensified delayed signal simultaneously. Most of the 100 MHz oscilloscopes have this facility. The availability of this feature is indicated by the ALT sweep display mode. To use this facility, depress ALT sweep display instead of main sweep.

B Ends A mode

Sometimes it is observed that when the delayed sweep to main sweep speed ratio is very high, the expanded display in the delayed sweep mode has somewhat reduced intensity. B Ends A mode can be used to increase the intensity of delayed sweep display by ending the main sweep at the minimum required point and increasing the display time for the delayed sweep. This happens because the slow main sweep runs for the full screen and there is very little time for the much faster delayed sweep.

Some oscilloscopes also have triggered delayed sweep facility. Operationally, it is similar to delayed sweep. In the delayed sweep mode, the delay time multiplier can be adjusted to smoothly move the intensified portion on the screen. In the triggered delayed sweep, the intensified portion jumps from one level transition to the next as the

adjustment is done. After selecting the desired transition level where you want to trigger the delayed sweep and after selecting a proper slope (+) for positive going and (-) for negative going transition - the delayed sweep is engaged. This mode gives a highly reduced display jitter as the sweep is triggered by a definite trigger signal level.

X-Y operation

In the X-Y mode, the horizontal axis of the oscilloscope also represents a voltage rather than time as is the case in the usual oscilloscope operation. The time base circuitry gets bypassed. The signal to be represented on the horizontal or X-axis is applied to the horizontal deflection input available on the front panel of the oscilloscope having X-Y mode feature.

CH3 input is the horizontal input. It has two selectable horizontal deflection factors of 100mV/div. and 1V/div. i.e. 100mV signal (in case of 100mV/div. selection) and 1V signal (in case of 1V/div. selection) will sweep the beam horizontally by one division. The other signal is applied to the vertical input (one of the two vertical inputs in a dual channel oscilloscope). The result is the desired X-Y display.

A major problem with this kind of X-Y mode of operation is that it offers an uncalibrated fixed sweep speed. This problem is, however, overcome in majority of modern dual channel scopes by letting one of the two vertical inputs to be used as a horizontal input in the X-Y mode. The oscilloscopes having this provision will have the letters 'X' and 'Y' written near the input connectors of the two channels to indicate X and Y inputs when we select the X-Y mode. Thus, both horizontal and vertical axes have variable calibrated deflection factors.

One can also notice that the vertical position control corresponding to vertical channel being used for X-input in X-Y mode can be used to deflect the X-Y display horizontally. X-Y operational mode has numerous applications like plotting transfer characteristics of devices and circuits, measuring phase difference between two given signals having same frequency, measuring an unknown frequency etc.

Z-axis input

The oscilloscope display has three components: the horizontal component (X-axis component), the vertical component (Y-axis component) and the beam intensity (Z-axis component). The intensity remains constant for a particular setting of the intensity control during normal operation. Most of the scopes have an external Z axis input located on the rear panel. A signal fed to this input can be used to modulate the intensity of the display. Use of this input in conjunction with vertical inputs has many interesting applications.

Simple AM Transmitter

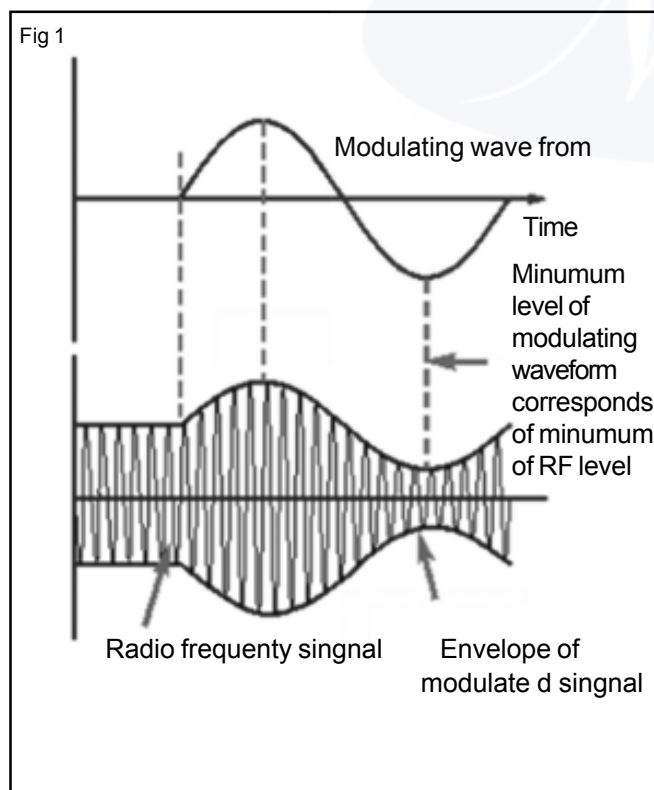
Objectives : At the end of this lesson you shall be able to

- what is Amplitude Modulation?
- simple AM Transmission

What is Amplitude Modulation

In order that a steady radio signal or "radio carrier" can carry information it must be changed or modulated in one way so that the information can be conveyed from one place to another. There are a number of ways in which a carrier can be modulated to carry a signal - often an audio signal and the most obvious way is to vary its amplitude.

Amplitude Modulation has been in use since the very earliest days of radio technology. The first recorded instance of its use was in 1901 when a signal was transmitted by a Canadian engineer named Reginald Fessenden. To achieve this, he used a continuous spark transmission and placed a carbon microphone in the antenna lead. The sound waves impacting on the microphone varied its resistance and in turn this varied the intensity of the transmission. Although very crude, signals were audible over a distance of a few hundred metres. The quality of the audio was not good particularly as a result of the continuous rasping sound caused by the spark used for the transmission.



Later, continuous sine wave signals could be generated and the audio quality was greatly improved. As a result, amplitude modulation, AM became the standard for voice transmissions.

Currently amplitude modulation is primarily used for broadcasting, but it is still used for some forms of two way radio communications. Its main radio communications use is for local aviation related VHF two way radio links. It is used for ground to air radio communications as well as two way radio links for ground staff as well.

Amplitude modulation basics

When an amplitude modulated signal is created, the amplitude of the signal is varied in line with the variations in intensity of the sound wave. In this way the overall amplitude or envelope of the carrier is modulated to carry the audio signal. Here the envelope of the carrier can be seen to change in line with the modulating signal.

Amplitude modulation, AM is the most straightforward way of modulating a signal. Demodulation, or the process where the radio frequency signal is converted into an audio frequency signal is also very simple. An amplitude modulation signal only requires a simple diode detector circuit. The circuit that is commonly used has a diode that rectifies the signal, only allowing the one half of the alternating radio frequency waveform through. A capacitor is used to remove the radio frequency parts of the signal, leaving the audio waveform. This can be fed into an amplifier after which it can be used to drive a loudspeaker. As the circuit used for demodulating AM is very cheap, it enables the cost of radio receivers for AM to be kept low.

Amplitude modulation advantages & disadvantages

Like any other system of modulation, amplitude modulation has several advantages and disadvantages. These mean that it is used in particular circumstances where its advantages can be used to good effect..

Advantages

Disadvantages

- It is simple to implement
- It can be demodulated using a circuit consisting of very few components

- AM receivers are very cheap as no specialised components are needed.
- An amplitude modulation signal is not efficient in terms of its power usage
- It is not efficient in terms of its use of bandwidth, requiring a bandwidth equal to twice that of the highest audio frequency
- An amplitude modulation signal is prone to high levels of noise because most noise is amplitude based and obviously AM detectors are sensitive to it.

In view of its characteristics advantages and disadvantages, amplitude modulation is being used less frequently. However it is still in widespread use for broadcasting on the long, medium and short wave bands as well as for a number of mobile or portable communications systems including some aircraft communications.

Amplitude modulation related signals

Amplitude modulation forms the basis of a number of forms of signal apart from the basic mode. These signal formats are typically generated by removing or suppressing the carrier, and then utilising the sidebands. These formats are defined in greater detail in further pages of this tutorial.

Description	Abbreviation	ITU Designation
Double sideband, full carrier i.e. basic	AMDSB full carrier	A3E
Single sideband reduced carrier	SSB reduced carrier	R3E
Single sideband full carrier	SSB full carrier	H3E
Single sideband suppressed carrier	SSBSC	J3E
Independent sideband	ISB	B8E
Vestigial sideband	VSF	C3F

While amplitude modulation is one of the simplest and easiest forms of signal modulation to implement, it is not the most efficient in terms of spectrum efficiency and power usage. As a result, the use of amplitude modulation is falling in preference to other analogue modes such as frequency modulation, and a variety of digital modulation formats. Yet despite this decrease, amplitude modulation is in such widespread use, especially for broadcasting, and many amplitude modulation signals can still be heard on the various long, medium and short wavebands where they will undoubtedly be heard for many years to come.

Simple AM Transmission

Transfer of information (speech, music, image, computer data etc.) by radio can be presented in its simplest form with block - diagram as on Fig 2. That is a transmission realized by amplitude - modulated signal. Since, in our example, the information being transferred is the sound, the first step of such transmission is converting the sound into electrical signal, this being accomplished by a microphone. The low - frequency (LF) voltage at microphone output (Fig 2), that represents the electrical "image" of the sound being transferred, is being taken into the transmitter. There, under the effect of LF signal, the procedure called amplitude modulation is being carried out, and on its output high - frequency (HF) voltage is generated, its amplitude changing according to the current LF signal value. HF voltage creates HF current in the antenna, thus generating electromagnetic field around it. This field spreads through the ambient space, being symbolically shown on Fig 2 with dashed circles. Traveling at the speed of light ($c=300\,000\text{ km/s}$), the electromagnetic field gets to the reception place, inducing the voltage in the reception antenna, as shown on Fig 2. This voltage has the same profile as the one on Fig 2, except it has much smaller amplitude. In the receiver, the amplification and detection are carried out first, resulting with the LF voltage on its output, that has the same profile as the one on Fig 2. This voltage is then transformed into sound by loudspeaker, that sound being exactly the same as the sound that acted upon the microphone. This, naturally, is the way it would be in ideal case. Back to reality, due to device imperfection as well as the influence of various disturbances, the sound being generated by the loudspeaker differs from the one that acts upon the microphone membrane. The block - diagram on Fig 2 (excluding the HF signal shape) is also applicable in case of radio transmission being carried out by frequency modulation. In that case frequency modulation is being carried out in the transmitter, under the effect of LF signal coming from the microphone, therefore HF signals on Fig 2 having constant amplitude, and their frequency being changed in accordance with the actual value of LF signal from the microphone. In fact, all types of radio transmission can be presented with Fig 2. First, the information being sent is always transformed into electrical signal through the appropriate converter. In telegraphy this converter is the pushbutton, in radiophony it's a microphone, in television engineering an image analysis cathode ray tube (CRT) etc. Then, with this "electrical image" of information, the modulation is being done. The modulated HF signal is being transferred into antenna and transmitted. On the reception place, the modulated signal from the reception antenna is being amplified and detected and then, again with the appropriate converter (pen recorder, loudspeaker, TV CRT etc.), the information is transformed back into its original form.

Fig 2

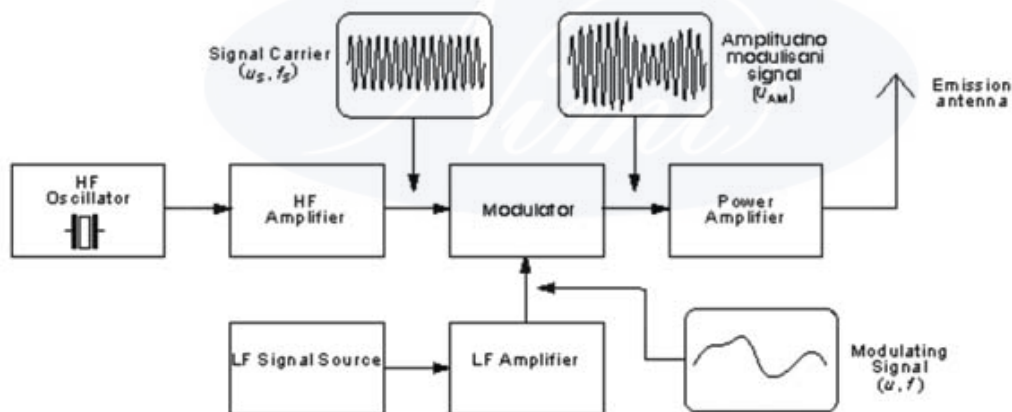


AM Transmitter

In order to better understand the way the radio transmitter works, block - diagram of a simple AM (amplitude modulated) signal transmitter is shown on Fig 3. The amplitude modulation is being performed in a stage called the modulator. Two signals are entering it: high frequency signal called the carrier (or the signal carrier), being created into the HF oscillator and amplified in the HF amplifier to

amplitude is being performed, therefore it is being called the modulating signal. The shape of the AM signal exiting the modulator is shown on Fig 3. From the point t_0 this voltage has the same shape as that on Fig 3. From the moment t_0 the amplitude of AM signal is being changed in accordance with the current value of the modulating signal, in such a way that the signal envelope (fictive line connecting the voltage peaks) has the same shape as the

Fig 3



the required signal level, and the low frequency (modulating) signal coming from the microphone or some other LF signal source (cassette player, record player, CD player etc.), being amplified in the LF amplifier. On modulator's output the amplitude modulated signal u_{AM} is acquired. This signal is then amplified in the power amplifier, and then led to the emission antenna.

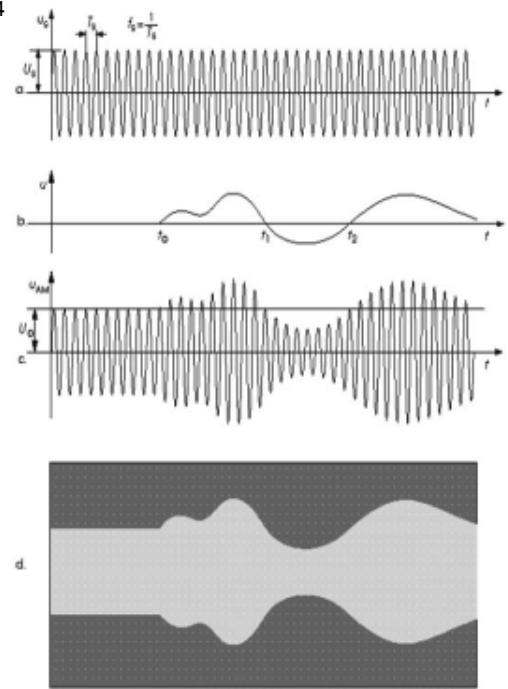
The shape and characteristics of the AM carrier, being taken from the HF amplifier into the modulator, are shown on Fig 4. As you can see, it is a HF voltage of constant amplitude u_s and frequency f_s . On Fig 4 the LF signal that appears at the input of the modulator at the moment t_0 is shown. With this signal the modulation of the carrier's

modulating signal. Let's take a look at a practical example. Let the LF signal on Fig 4 be, say, an electrical image of the tone being created by some musical instrument, and that the time gap between the points t_0 and t_2 is 1 ms. Suppose

that carrier frequency is $f_s = 1$ MHz (approximately the frequency of radio Kladovo, exact value is 999 kHz). In that case, in period from t_0 till t_2 signals us on Fig 4 and u_{AM} on 2.3-c should make a thousand oscillations and not just eighteen, as shown in the picture. Then it is clear that it isn't possible to draw a realistic picture, since all the lines would connect into a dark spot. The true picture

of AM signal from this example is given on Fig 4. That is the picture that appears on screen of the oscilloscope, connected on the output of the modulator: light coloured lines representing the AM signal have interconnected, since they are thicker than the gap between them. Block - diagram on Fig 3 is a simplified schematic of an AM transmitter. In reality there are some additional stages in professional transmitters that provide the necessary work stability, transmitter power supply, cooling for certain stages etc. For simple use, however, even simpler block diagrams exist, making the completion of an ordinary AM transmitter possible with just a few electronic components.

Fig 4



Crystal AM receiver

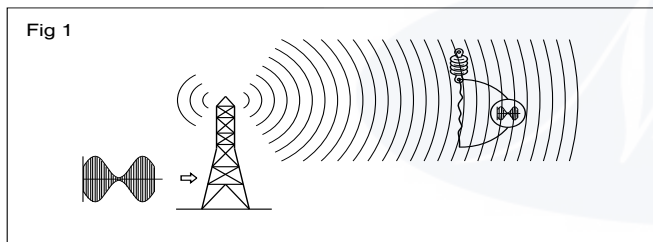
Objectives: At the end of this lesson you shall be able to

- explain the basic principle of radio reception
- explain the basic principle of selecting one of several signals available in space
- name the two main frequency bands used for commercial AM radio broadcasting
- sketch the schematic of a crystal receiver and explain the working of the circuit
- explain the terms- selectivity, sensitivity, fidelity and signal to noise ratio
- list the limitations in crystal receivers and explain how the limitations of a crystal receiver are over-come in TRF receivers
- sketch the schematic of a TRF receiver using integrated circuits.

Principle of Radio receivers

The modulated wave emitted from a transmitting antenna in the form of electromagnetic waves(energy) travels in free space at the speed of light(3×10^8 metres/sec). The distance these electromagnetic waves travel depends upon the chosen type of transmitting antenna and the transmitted power.

As shown in Fig 1, if a wire is suspended above the ground in the region where the electromagnetic waves are travelling, the passing electromagnetic waves induce a small voltage in the wire.

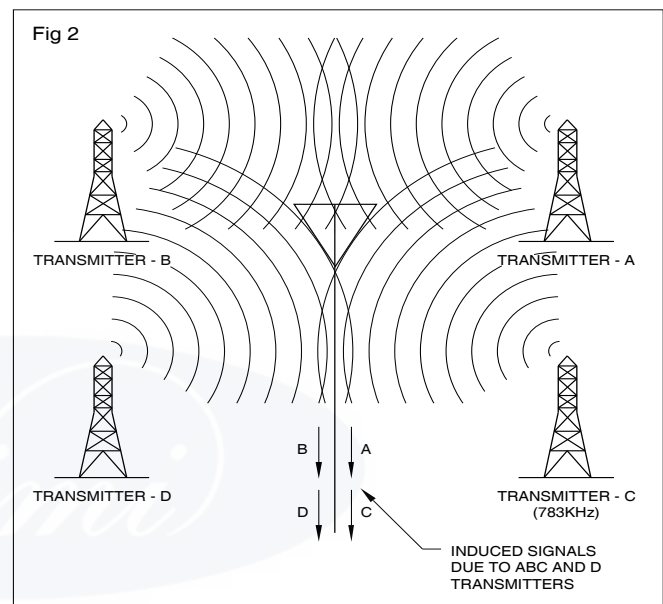


Although the induced voltage in the suspended wire is very small, of the order of microvolts, the wave-form of the induced voltage is an exact replica of the signal transmitted by the transmitting antenna. Thus the suspended wire acts as a receiving antenna.

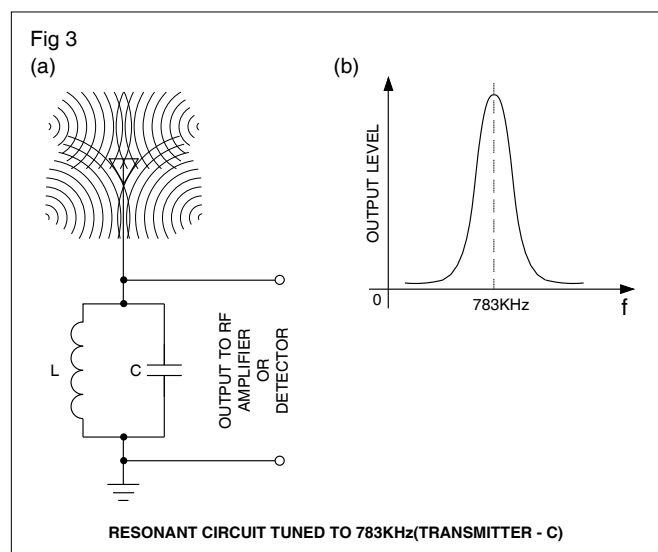
If this received weak signal, at the receiving antenna is processed, the information (voice and music etc.) which modulated the carrier at the transmitter can be reproduced.

Selecting required Electromagnetic wave (signal)

Since several transmitting antennas of several stations transmit information (voice and music etc) simultaneously into air, the electromagnetic waves of several stations simultaneously exist in free space. All these electromagnetic waves corresponding to several stations intercept(cut) the receiving antenna at the same time. Each of these simultaneously induce voltages in the receiving antenna as shown in Fig 2. Hence, the receiving antenna or the aerial will have signal voltages corresponding to several transmitting stations as shown in Fig 2.



The first job of a radio receiver is to select a particular station signal in which we are interested and reject the rest of signals. This can be done easily using a tuned circuit as shown in Fig 3a.



As already discussed in previous units, the frequency response of the parallel tuned circuit at Fig 3a will be as shown in Fig 3b. How good is this parallel tuned circuit, in selecting a particular frequency signal is termed as the *selectivity* of the circuit. This selectivity of the tuned circuit in a receiver decides whether or not the receiver suffers from interference of unwanted stations.

AM receiver and frequency bands

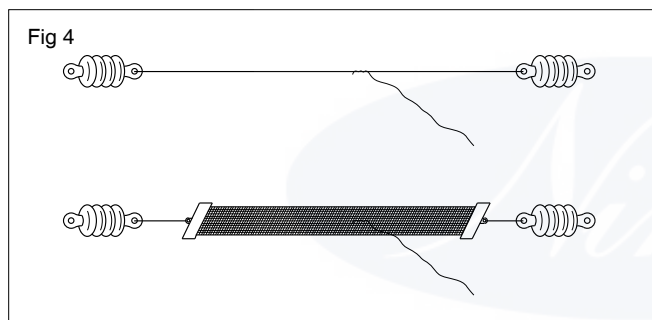
At present AM broadcasting is restricted to the following frequency bands;

Medium Wave(MW) band	535 KHz to 1605 KHz
Short Wave(SW) band	3 MHz to 26 MHz

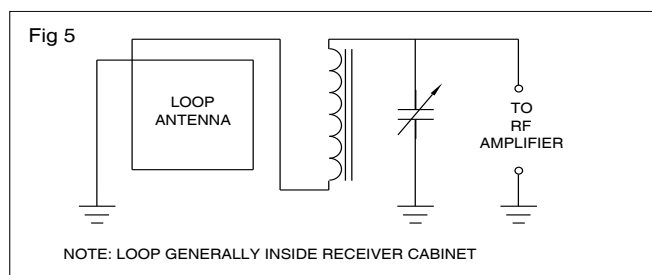
Broadcasting stations, transmitting amplitude modulated (AM) MW band signals relay on ground waves for propagation. The SW band AM transmitters on the other hand relay both on ground waves and sky waves for propagation. Hence, the distance covered by MW AM stations is much less than that of SW AM stations.

The receiving antenna of AM receivers

In earlier days a lengthy copper wire or a lengthy copper tape made of thin copper wire mesh as shown in Fig 4 were used as receiving antennas for AM receivers.

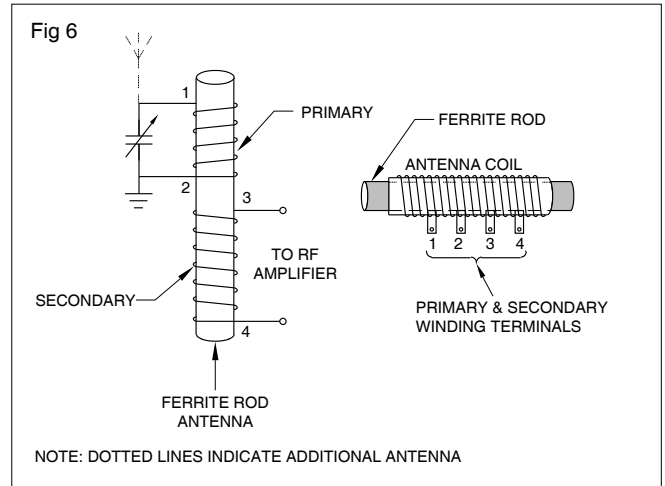


Due to the large space occupied by these antennas, and for use with portable AM receivers, a loop antenna coupled to an antenna transformer as shown in Fig 5 is used in present day AM receivers.



Recent innovation, is the ferrite rod antenna as shown in Fig 6. This antenna, with ferrite, having an extremely high permeability has excellent pick-up/receiving characteristics.

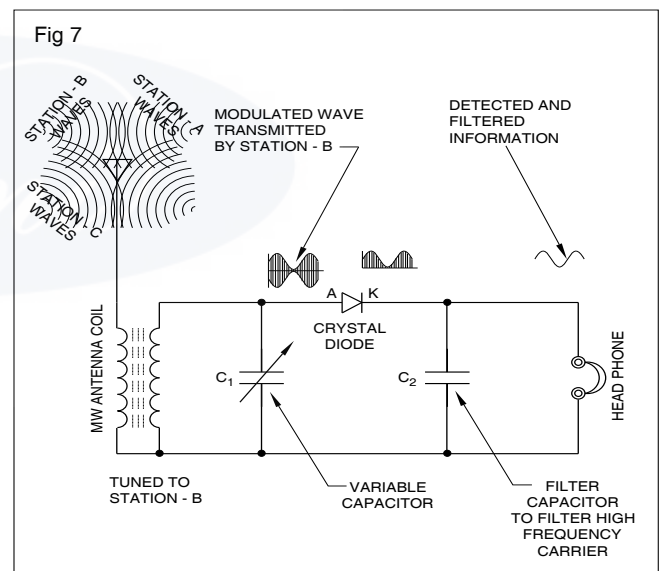
The antenna transformer used to couple the selected signal to the next stage is often referred to as **antenna coil** rather than antenna transformer. The primary winding of the antenna coil(antenna transformer) isolates the antenna side from the tuned circuit and provides proper impedance matching to the antenna.



Generally ferrite core antenna coil is wound on insulating material such as paper or plastic. The gauge of wire used and the number of turns used for the antenna coils designed to receive MW band stations and the SW band stations differ.

Basic crystal radio receiver

The simplest form of a radio receiver is the crystal receiver shown in Fig 7.



The basic crystal AM receiver consists of an L-C resonant circuit (antenna coil and a variable capacitor), a crystal diode, a filter capacitor and a headphone.

Station Selection stage

In a crystal receiver, the desired station frequency is selected by the L-C parallel resonant circuit whose resonance frequency is altered by tuning the variable capacitor. Once the parallel L-C circuit is tuned to a desired station frequency, the voltage across the parallel tuned circuit will be maximum for that frequency, thus giving a modulated voltage at the anode of the rectifier diode as shown in Fig 7.

Detection or demodulation stage

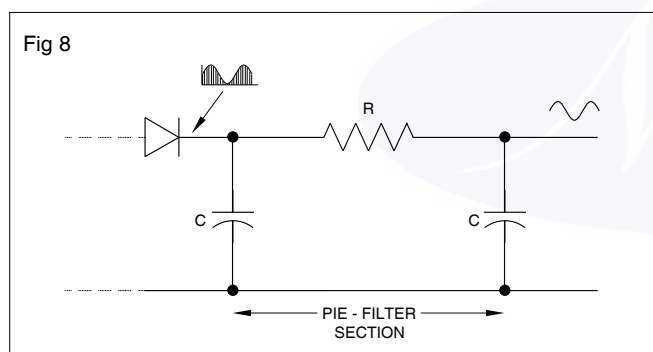
The modulated wave available at the input (Anode) of the diode has both positive and negative envelopes of the modulating signal and the carrier frequency. Recall that the information transmitted is in the envelope of the modulated RF carrier wave and not in the carrier wave.

In Fig 7, the crystal diode used rectifies the modulated wave to get the positive side of the modulated wave (negative side may also be chosen as the information contained is the same). This process is called **Demodulation or Detection**.

A germanium crystal diode is used for detection because, the amplitude of the received signal is very weak. If a silicon diode was used, it would need a minimum 700 mV to conduct, and hence, it fails to detect low voltage levels.

Filtering the carrier stage

The rectified modulated wave contains the RF carrier frequency component in addition to the information as its envelop. If this was fed to the head phone directly, it would result in excessive noise along with the information (envelop). Hence, this RF carrier component in the detected signal has to be eliminated before feeding the detected signal to a head phone. This can be easily done by filtering, using a capacitor. Usually a pie-filter as shown in Fig 8 is used to trap the RF carrier completely.

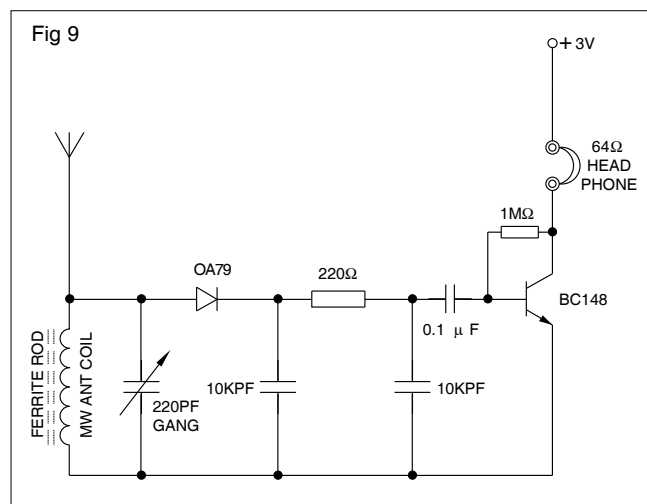


Note that, the basic crystal receiver shown in Fig 7 uses a single capacitor instead of a pie-filter. This is basically to avoid the loss across the resistor in a pie-filter. If suitable amplification of the received modulated signal is carried out, then a pie-filter can be used as is generally done in commercially available AM radio receivers.

Audio stage

The detected and filtered audio signal can be connected to a high impedance (>32W) headphone which will reproduce the information in the form of sound waves. Most radio receivers, however, amplify the detected and filtered signals to suitable levels using audio amplifiers and then fed to a loud speaker.

Fig 9 shows a crystal receiver with a pie-filter for improved Rf filtering and a transistor amplifier for improving the audio output level. Since an amplifier is used, the circuit now requires DC power supply.



RECEIVER CHARACTERISTICS

Selectivity

One of the problems in a crystal receiver is its inability to sharply select the required station. This station selection performance of a radio receiver depends on the characteristics of the tuned LC circuit which has to accept only the desired frequency, while rejecting the rest. This characteristic is termed as *selectivity* of the receiver.

Selectivity is the ability of a receiver to tune-in or receive a desired channel frequency while rejecting other undesired channel frequencies.

It is important to note that a very high selectivity does not necessarily make a better receiver. This is because, a radio broadcast signal consists of the carrier and its side bands which are the sum and difference of frequencies produced by the modulating signal. For example, if a 5 KHz audio tone modulates a carrier of 1000 KHz the frequencies produced after modulation are,

Carrier : 1000 KHz

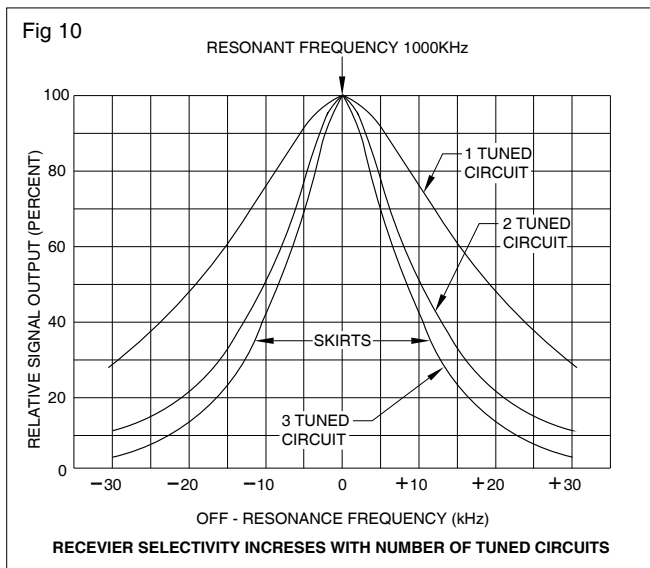
Upper side band : $1000 \text{ KHz} + 5 \text{ KHz} = 1005 \text{ KHz}$

Lower side band : $1000 \text{ KHz} - 5 \text{ KHz} = 995 \text{ KHz}$

Thus for the receiver to properly reproduce the modulating signal, the selecting section of the receiver has to pass a band of frequencies from 995 KHz to 1005 KHz. This means that the selecting section of the receiver should have a band width of 10 KHz. A receiver more selective than this, cannot reproduce the entire modulating information.

Hence, the selection section of the receiver must pass the 10 KHz band around the 1000 KHz carrier, but it must reject frequencies outside this band which otherwise may ride through the receiver and heard as *background sounds*.

Selectivity of a receiver is determined by the bandwidth of the tuned circuit in the receiver. This in turn depends upon the Q of the tuned circuit. Hence, if the required selectivity cannot be obtained by a single tuned circuit, (due to the limitation in the value of achievable Q) the required selectivity can be obtained by increasing the number of tuned circuits used in a receiver as shown in Fig 10.



Another problem in crystal receivers is that it is able to receive only the nearby transmitting stations and not all the stations in the MW band. This is because the crystal receiver is not sensitive enough to pick up weak signals coming from far off transmitting stations.

Sensitivity of a receiver is the ability of the receiver to respond satisfactorily to weak signal voltages. In other words, sensitivity is the minimum signal input voltage required to produce a specified output.

Sensitivity of a receiver is measured in terms of the voltage that must be induced in the antenna by electromagnetic signals to develop a standard audio output from the amplifier. This standard output is arbitrarily chosen as 50 mWatts.

The sensitivity of broadcast receivers is not more than 100 microvolts. Under most favourable conditions, specially designed communication receivers have the highest sensitivity of the order of 1 microvolt.

The lower the numerical figure indicated for sensitivity, the better is the receiver. For example, if receiver-A is said to have a sensitivity of 100 microvolts, it means a minimum of 100 μ volt is necessary for this receiver to function satisfactorily. If another receiver-B is said to have a sensitivity of 25 microvolts, it means the 25 microvolts itself is sufficient for this receiver to function satisfactorily. Hence, receiver-B is more sensitive compared to receiver-A.

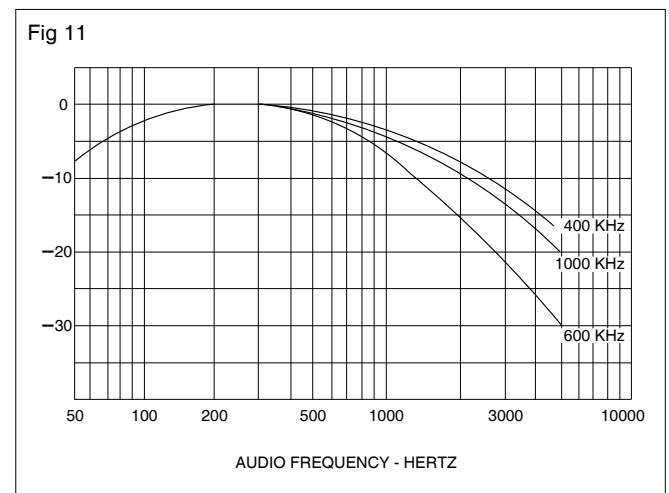
The third problem with crystal receivers is the poor quality of the audio signal heard from the headphones. This is because of the low signal level at the input and the poor processing of the detected audio signal before it is fed to a loudspeaker.

Fidelity is the term used to indicate the accuracy of reproduction at the output of the radio receiver. Fidelity of a receiver is the ability of the receiver to truly re-produce the audio signals transmitted.

Usually it is found that the sound produced by one radio is different from another radio. This is because the radio receiver may not be in a position to reproduce all the audio frequency transmitted by the transmitters.

The fidelity of the receiver depends on various factor such as, carrier frequency alignment of the receiver, frequency response of the A.F amplifier and the loudspeaker.

A graph showing the fidelity characteristics of a typical radio receiver is shown in Fig 11. Ideally this graph should be flat.



Another important parameter associated with radio receivers, that define the quality of the radio receiver, is the signal to noise ratio.

Signal to noise ratio: It is the ratio of signal power at the output to the noise power at the output at the given R.F frequency. Usually the ratio is found out at various frequencies in the given band and a curve is drawn. The ratio is always expected to be very high. In other words the higher the signal-to-noise ratio, the better is the radio receiver.

Limitations of the crystal receiver

Crystal receivers is important only during the initial stages of understanding the basics behind radio receivers. Commercially crystal receivers are not of any use because they suffer from four major disadvantages. They are as follows;

- Low selectivity - because it has a single tuned circuit. Hence such a receiver will have interference from the transmitting stations operating with carrier frequency close to the wanted station carrier frequency.
- Low sensitivity - because it does not have any RF amplification stages to amplify the received modulated signal of very low levels.
- Low fidelity - because it does not have sufficient RF and AF amplifiers and necessary audio signal processing circuit.
- Low signal-to-noise ratio - because the signal level is not suitably increased nor is there a circuit to suppress noise.

Hence crystal receivers can be used only in places which are very close to the transmitting stations where the modulated signals are strong. Also crystal receivers are used where sound quality is not important.

Tuned Radio Frequency receiver (TRF receiver)

Limitations in the crystal receiver can be largely overcome by what are known as TRF receivers. This type of receiver essentially consists of a chain of radio frequency amplifiers which improves the selectivity and the sensitivity of the radio receiver. A typical TRF receiver is shown in Fig 12.

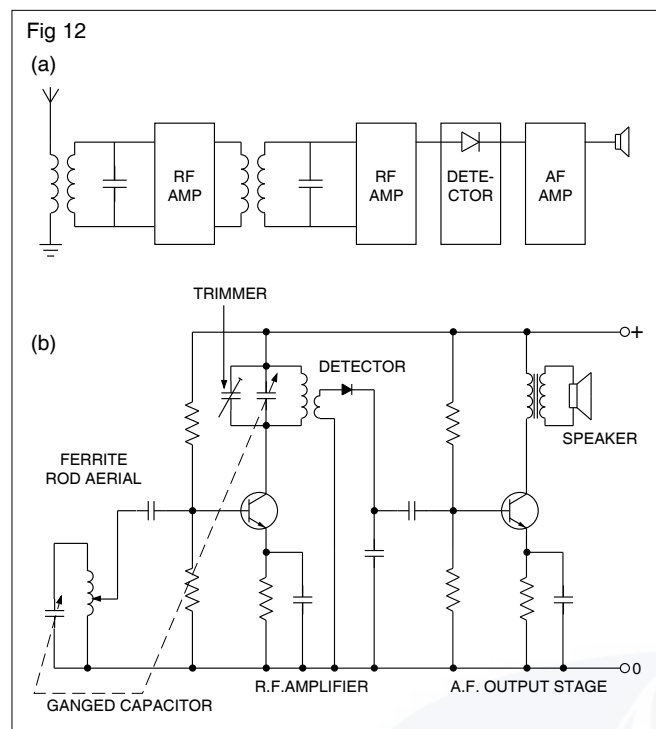


Fig 12a shows a TRF receiver with tuned RF amplifiers. Fig 12b shows the circuit of a TRF receiver with two tuned circuit. Since two tuned circuits are used,

- the selectivity is improved over that of simple crystal receiver
- the sensitivity is also improved because we have introduced a stage of RF amplification before the detector circuit.

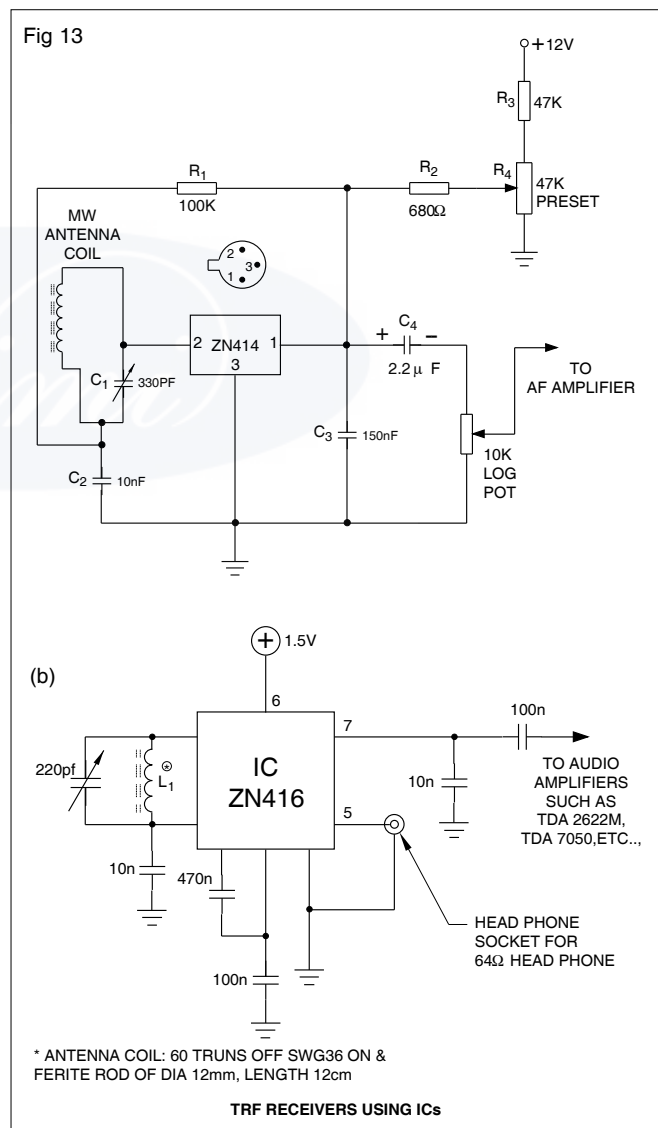
Although theoretically it is possible to further improve the sensitivity of the receiver by adding additional RF amplifier stages, in practice several problems and difficulties arise if the number of RF amplifier stages is more. Some of these difficulties are listed below;

- Because of the instability problem, the number of RF stages that can be used is limited. It is practically impossible to build a very high gain R.F amplifier merely by connecting several stages together. Each stage is obviously operating at the same frequency and it is very easy to have positive feedback from the latter stages to the input. Even though the voltage feedback may be very small, because of the high gain of the amplifier stages, conditions for self-oscillation are almost always present. Therefore, the amplifier may turn around and work as an oscillator.
- The other problem is to do with tracking. Every stage has its own tuned circuit and it is necessary to vary the resonant frequency of each tuned circuit in such a way

that all tuned circuits have the same resonant frequency. Apart from the practical difficulty of obtaining **ganged capacitors** with many sections, the inevitable stray capacitances in the various parts of the circuit upset the matching of each stage.

- The next problem has to do with the fidelity of the final audio output. Although addition of each tuned circuit improves the selectivity, and results in excellent sensitivity, the overall response becomes very sharp. It is as though a single tuned circuit with an extremely high Q value is being used. In order to receive the transmitted program information, a receiver which can respond, not only to the carrier, but also to the sidebands is required. As the receiver becomes more and more selective, the sidebands gets eliminated increasingly. Therefore, the audio frequency response becomes restricted.

Although the above drawbacks severely limit the use of TRF-receivers these receivers are used in some applications



in which the radio receiver is expected to receive only a few stations and is used sparingly.

The advantages of TRF radio receivers over crystal receivers are;

- Better selectivity
- Better sensitivity
- Improved signal to noise ratio.

TRF receivers using integrated circuits

Integrated circuits such as ZN414 and ZN416 (by Ferranti) are available commercially for constructing TRF receivers.

A simple circuit of a TRF receiver using ZN414 and ZN416 is shown in Fig 13.

The electromagnetic/radio waves are concentrated within the 6"x3/8" ferrite rod around which is wound 60 turns of thin insulated wire. The type of wire is not important, but the number of turns should be close together and wound in the same direction. The variable capacitor should be of solid dielectric type. The layout for the circuit is not critical, but the leads should be kept as short as possible. The output could be fed to a high quality, low impedance headphone using any speaker transformer or a power amplifier can be driven using the output. This receiver is best suitable for micro-sized radio receivers fitted right on the headphone itself.

For further details about ZN series of TRF receiver ICs, refer application notes of Ferranti.



Frequency modulation and FM receivers

Objectives: At the end of this lesson you shall be able to

- explain the terms FM, Modulation index 'm' and frequency deviation d in FM
- distinguish between AM and FM
- state the permissible maximum frequency swing in FM as per international standards
- state the allocated FM broadcast band for commercial FM transmission
- list the consequences of VHF and UHF transmission
- name the two methods of developing FM
- explain the FM generation using varicap diode through a circuit diagram
- name the blocks of a typical FM receiver
- explain the functions of *limiter circuit* and *discriminator circuit* in FM receiver
- explain the advantage of a *ratio detector* over discriminator circuit
- list a few other method of detecting FM signals.

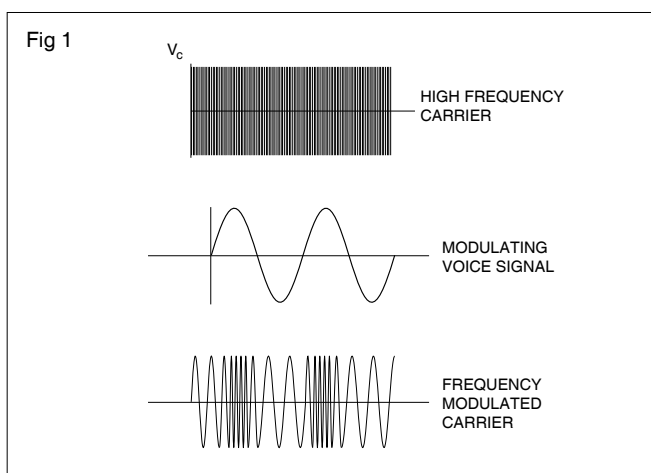
FREQUENCY MODULATION

Frequency modulation, or FM, is an alternative to amplitude modulation. Frequency modulation has some advantages that make it attractive for commercial broadcasting and two-way radio work.

One problem with AM is its sensitivity to noise. Lightning, automotive ignition, and sparking electric circuits all produce electric interference. This interference is spread over a wide frequency range. It is not easy to prevent such interference from reaching detector in an AM receiver. But, an FM receiver can be made totally insensitive to such noise interference. This noise-free performance of FM receivers has made them popular for several applications such as commercial radio reception, wireless communication sets etc.

Principle of Frequency Modulation(FM)

In frequency modulation the frequency of the RF carrier wave is varied in rhythm with the audio frequency oscillations as shown in Fig 1.



A variation in frequency is a change in the number of cycles per second. So, an increase in frequency is indicated in Fig 1 as a clustering or bunching of the RF oscillations (along the time scale). Similarly a decrease in frequency will appear as spreading against the time scale.

How far the frequency will vary (in other words, how thick will be the bunching in the graph) depends upon the amplitude of the AF oscillation (i.e., loudness of sound). This is analogous to the amplitude peak in the case of amplitude modulation. Also, the frequency of the AF oscillation (i.e., pitch of sound) will determine the number of such bunching and spreading per second.

To sum up, in the case of frequency modulation,

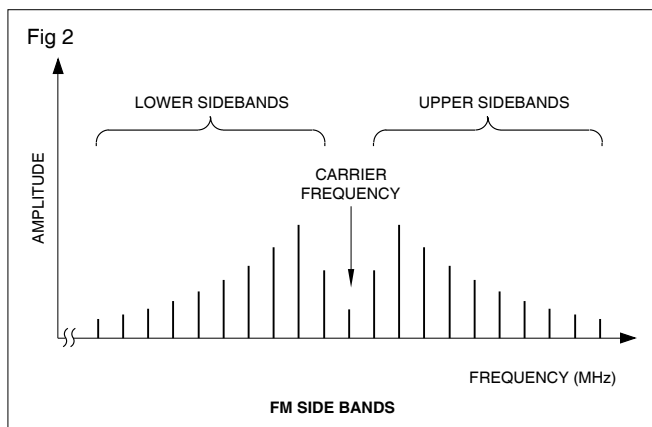
- 1 The loudness of the sound to be broadcast determines the amount of change in the frequency of the RF carrier wave (the thickness/density of the bunching in Fig 1).
- 2 The pitch, i.e., frequency of the broadcast sound determines the number of bunching i.e., the number of frequency peaks per second in the modulated wave.

Modulation Index m in Frequency Modulation

In amplitude modulation (AM), it is possible to vary the amplitude of the carrier wave (RF) between zero and twice its peak value for 100% modulation; but in Frequency modulation (FM), the variation of the carrier frequency between zero to twice the frequency value is neither possible nor permissible. By international agreement, the limit of frequency swing has been fixed at maximum intensity of the broadcast sound, the carrier frequency will vary within the limit of 150 KCs only (± 75 KC/s around the carrier central frequency).

As in the case of AM, Frequency modulated wave is also composed of carrier wave and side bands. Unlike the simple composition of 2 side frequencies besides the carrier in AM with one pure modulating frequency, in FM, the number of side frequencies or side bands for each modulating AF appear to be unlimited as shown in Fig 2.

In FM although the number of side bands appear to be unlimited, the strength of the higher side bands becomes increasingly so feeble that beyond a certain number the higher sidebands may be omitted without impairing the quality of the transmitted signal.



The side band frequencies are:

$$1\text{st side band} = F_r \pm f_m$$

$$2\text{nd side band} = F_r \pm 2f_m$$

$$3\text{rd side band} = F_r \pm 3f_m \quad \text{and so on...}$$

Where,

F_r is the Carrier frequency (RF)

f_m is the Modulating frequency (AF)

The limit is approximately at $1.5 \times m_{\text{max}}$; Where, m_{max} is the maximum value of the modulation index. Modulation index m denotes the ratio of frequency swing (df_r) to the modulating frequency (f_m).

Thus,

$$m = \frac{\text{Frequency swing}}{\text{modulating frequency}} = \frac{df_r}{f_m}$$

$$= \frac{75 \text{ Kc/s}}{f_m} \quad \text{at its maximum value.}$$

If the transmission is to be high fidelity, all the significant side bands are to be retained during transmission and reception. Therefore the total channel width or band-width for Hi-Fi FM transmission reception should be $150 \text{ Kc/s} \times 1.5 = 225 \text{ Kc/s}$.

FM Broadcast Band

In view of the large bandwidth occupied by the FM transmitters it is not possible to range these transmitters in the MW or even SW bands. By international agreement the frequency band of 86.5MHz to 100 MHz (3.4m to 3 m) is reserved for FM broadcasting.

Besides the special FM technique, attention should also be devoted to specific very-short-wave (VHF) or ultra-short-wave (UHF) transmission problems.

VHF or UHF transmission: With wave lengths smaller than 5 meters, the transmission is principally by direct route or by line-of-sight propagation. Because of this, the two main consequences are,

- 1) A great part of the radiated energy disappears in the infinite space as bending by ionosphere does not takes place.
- 2) The distance of transmission of the transmitter is limited by the rounding of the earth.

Theoretically it means that reception of FM signals is possible only when transmitter and receiver aerials can see each other. But, practically the distance of FM transmission is slightly greater than the theoretical distance. This means, the communication range for FM transmission is generally limited to around 100Km. The possible distance (d) which can be bridged, may be roughly calculated using the thumb rule formula as follows;

$$d = 4 (\sqrt{H} + \sqrt{h})$$

Where d = the distance in Km from transmitter to receiver.

H = height of transmitting antenna .

h = height of the receiving antenna in meters.

Principle Frequency Modulation Circuits

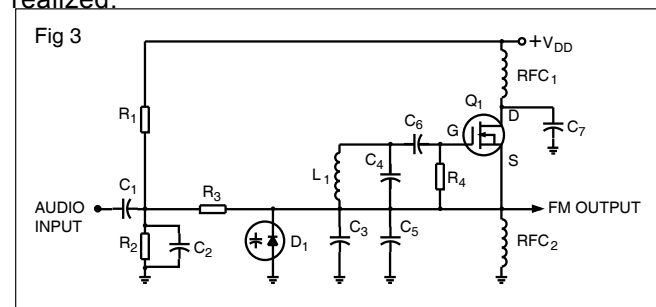
The two most popular methods of developing Frequency modulated signal are by,

- **direct method** in which an oscillator is modulated directly by the audio signal.
- **indirect method** in which Phase modulation is used to vary the carrier to create an FM signal. In this the oscillator used will generally be a crystal controlled oscillator.

Phase modulation is another type of modulation in which the phase angle of the carrier is varied in accordance with the variations in the amplitude of the modulating AF signal. Note that phase modulation also creates frequency modulation. For further details on phase modulation refer reference books listed at the end of this book.

The simplest method of producing an FM signal is by direct method, wherein the modulating signal vary the value of L or C in an oscillator causing the frequency to vary. The basic concept is by having a capacitor microphone placed directly across the tuning capacitor of a Hartley oscillator. The Hartley oscillator frequency is determined by the total L and C in the tank circuit. When the sound waves strike the movable plate of the capacitor microphone, the capacitance will increase or decrease in accordance with the loudness of the sound waves. More capacitance means less frequency; less capacitance means more frequency, thus the output of the oscillator is a frequency modulated (FM) wave. In refined systems for producing FM involves use of a varicap diode (varector diode) to vary the frequency of the tank circuit

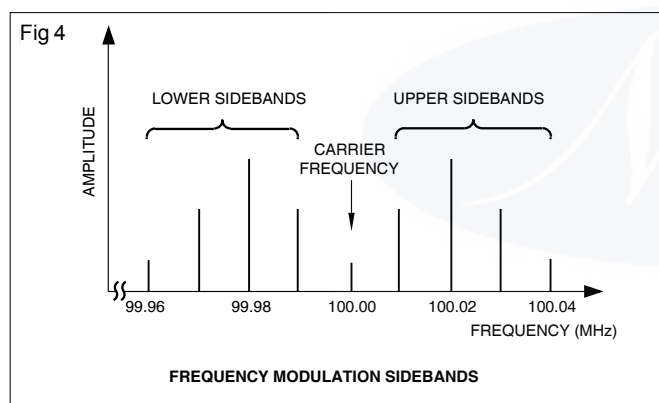
Figure 3 shows how frequency modulation can be realized.



Transistor Q_1 and its associated parts make up a series-tuned Colpitts oscillator. Capacitor C_3 and coil L_1 have the greatest effect in determining the frequency of oscillation. Diode d_1 is a variable capacitance (varicap) diode. It is connected in parallel with C_3 . This means that as the capacitance of D_1 changes, so will the resonant frequency of the tank circuit. Resistors R_1 and R_2 form a voltage divider to bias the varicap diode. Some positive voltage (a portion of V_{DD}) is applied to the cathode of D_1 is in reverse bias.

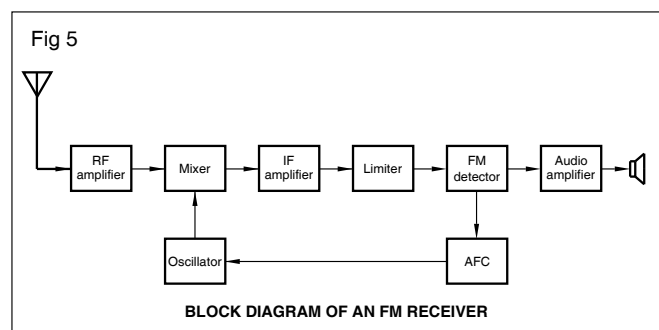
A varicap diode uses its depletion region as the dielectric. More reverse bias means a wider depletion region and less capacitance. Therefore, as an audio signal goes positive, D_1 should reduce its capacitance. This will shift the frequency of the oscillator up. A negative-going audio input will reduce the reverse bias across the diode. This will increase its capacitance and shift the oscillator to a lower frequency. Thus, the audio signal is modulating the frequency of the oscillator.

Suppose the oscillator frequency was 100 MHz (commercial FM station frequency), and this frequency is being modulated with a steady 10 KHz (0.01 MHz) audio signal, the sidebands produced are spaced 10 KHz apart as shown in Fig 4. These side bands appear above and below the carrier frequency as shown in Fig 4. From the frequency domain graph shown in Fig 4 it should be obvious that an FM signal will require more bandwidth than an AM signal.

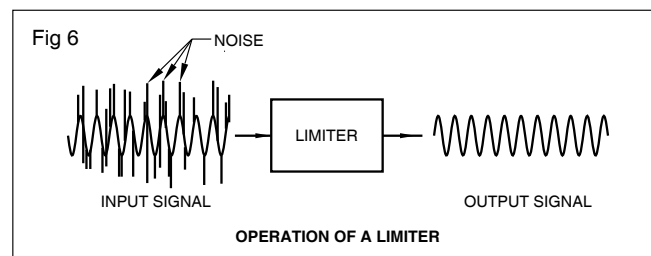


FM RECEIVERS

The block diagram for an FM superheterodyne receiver is shown in Fig 5.



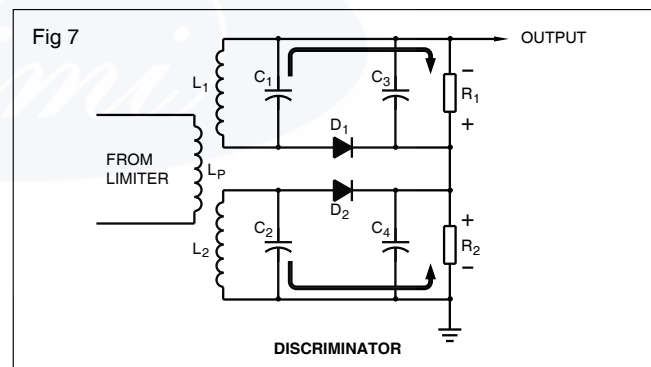
A typical FM receiver block diagram is quite similar to that of AM receiver. However, notice that a *limiter* stage appears between the IF stage and the detector stage. This is one way an FM receiver can reject noise. Fig 6 shows what happens in a limiter stage.



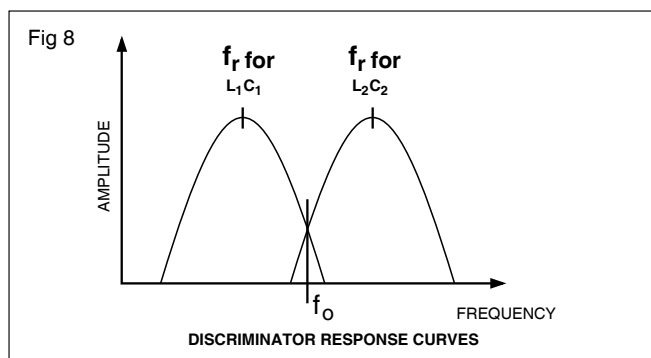
In the limiter shown in Fig 6, the input signal is very noisy. The output signal is noise-free. By limiting or by amplitude clipping, all the noise spikes have been eliminated. Some FM receivers use two stages of limiting to eliminate most noise interference.

Limiting cannot be used in an AM receiver because the amplitude variations carry the information to the detector. In FM reception, the frequency variations contain the information. Amplitude clipping in a FM receiver will remove just the noise but does not remove the information.

Detection in FM is more complicated than in AM. Since FM contains several side bands above and below the carrier, a single nonlinear detector (such as a diode) will not demodulate the signal. To detect a FM signal needs a double tuned *discriminator* circuit as shown in Fig 7.



The discriminator circuit shown in Fig 7 can serve as an FM detector. The discriminator works by having two resonant points. One is above the carrier frequency, and one is below the carrier frequency. The frequency response curves of this discriminator is shown in Fig 8.



In the frequency response curves for the discriminator circuit, f_o represents the correct point on the curves for the carrier. In a FM receiver, the station's carrier frequency will be heterodyned to f_o . This represents a frequency of 10.7 MHz for broadcast FM receivers. The heterodyning process allows one discriminator circuit to demodulate any signal over the entire commercial FM band.

In the discriminator circuit shown in Fig 7, when the carrier is unmodulated, D_1 and D_2 will conduct an equal amount. This is because the circuit is operating where the frequency response curves cross. The amplitude is equal for both tuned circuits at this point. The current through R_1 will be equal to the current through R_2 . If R_1 are equal in resistance, the voltage drops will also be equal. Since the two voltages are series-opposing, the output voltage will be zero. This means, when the carrier is at rest (no modulation), the discriminator output is zero.

Suppose the carrier shifts higher in frequency because of modulation, this will increase the amplitude of the signal in L_2C_2 and decrease the amplitude in L_1C_1 . Hence there will be more voltage across R_2 and less across R_1 . Thus, the output of the discriminator goes positive.

On the other hand when the carrier shifts below f_o , the signal is closer to the resonant point of L_1C_1 . More voltage will drop across R_1 , and less across R_2 . Hence, the output goes negative.

The output from the discriminator circuit will be,

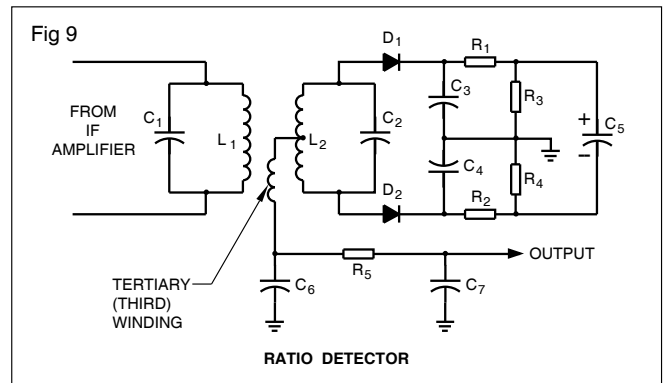
- zero when the carrier is at rest,
- positive when the carrier moves higher in frequency, and
- negative when the carrier moves lower in frequency.

Thus the output of the discriminator is a function of the carrier frequency.

Automatic Frequency control

The output of the discriminator can also be used to correct any drift in the receiver oscillator frequency. As shown in Fig 5, FM detector feeds a signal to the audio amplifier and to a stage marked AFC. AFC stand for *automatic frequency control*. If the oscillator output frequency happens to change for some reason, then, f_o will not equal exactly 10.7 MHz. There will be a steady dc output voltage from the discriminator. This dc voltage can be used as a control voltage to change the oscillator frequency automatically and set it back to 10.7 MHz.

The discriminator circuits discussed above work fairly well; but they are sensitive to amplitude changes. This is why one or two limiters are needed for noise free reception. An improved method of FM detection is by using the *ratio detector* instead of the discriminator. Ratio detectors are not very sensitive to the amplitude of the signal. This makes it possible to build receivers without limiters and still provide good noise rejection. Fig 9 shows a typical *ratio detector* circuit.



The design of ratio detector is based on the idea of dividing a single voltage into a ratio. This ratio is equal to the ratio of the voltages on either side of L_2 . With frequency modulation, the ratio shifts and an audio output signal is available at the center tap of L_2 . Since the circuit is ratio-sensitive, the input signals amplitude may vary over a wide range without causing any change in output. This makes detector insensitive to amplitude variations such as noise.

There are quite a few other FM detector circuits used in FM receivers. Some of the more popular ones are,

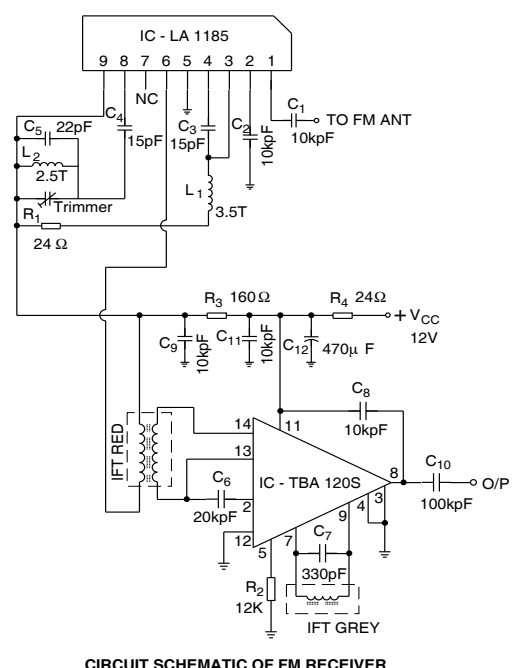
- the quadrature detector,
- the phase-locked-loop detector, and
- the pulse-width detector.

These circuits are generally used in conjunction with integrated circuits (IC's). They usually have the advantage of requiring no alignment or only one adjustment whereas alignment for discriminators and ratio detectors is more time-consuming.

The circuit diagram of a simple portable FM receiver using ICs is shown in Fig 10.

For further reading on FM transmitters and receivers, refer reference books listed at the end of this book.

Fig 10



Digital IC families and their operational characteristics

Objectives: At the end of this lesson you shall be able to

- explain the basic terms related to digital IC gates
- recognize the different types of packages of IC(s) used in the digital IC
- list different levels of integration used in fabrication of digital IC
- differentiate logic families and their characteristics
- compare the TTL and CMOS families
- explain digital IC numbering system.

INTRODUCTION

A digital system is a combination of devices designed to Process information that are represented in digital form. Example of a few most popular digital systems are,

- Digital computers
- Calculators,
- Digital audio and video equipment
- Telephone system etc.,

Digital Telephony is probably the world's largest digital system.

In electronic circuits, signals are represented in voltage or current. In these circuits, the signal representation will have a number of voltage or current levels.

In analog signals, the transition from one level to another is usually smooth rather than sudden.

Digital signals on the other hand can have only two discrete states. These states can be called as,

- **ON state:** A state at which a predefined voltage is present. For example, the level could be, +5 Volts, +10Volts and it is also represented as high, one, etc.
- **OFF state:** A state at which a predefined voltage is other than the ON state voltage is present. For example, the level could be, 0 Volts, -5 Volts and it is also represented as low, zero etc.,

The discrete levels in digital signals are technically referred to as logic levels. Generally, the ON state described above is referred as the LOGIC 1 state and the OFF state as the LOGIC 0 state. It is very essential to note that, in digital signal representation, no state exists in between the logic-0 and logic-1 state.

For example, if we say Logic-0 corresponds to 0 volts and Logic-1 corresponds to 1 volt. In such a digital system, voltage levels of 2v, 3v, 4v etc., have no meaning (further details are discussed in lessons that follows).

Because the transition time between ON to OFF state or vice versa is abrupt in digital signals, analysis of digital systems varies from that of pure analog systems such as amplifiers etc.,

Compared to analog circuits, digital circuits contains less number of discrete components such as resistors, capacitors etc., This is mainly for the reason that the Integrated circuit(IC) technology has advanced so much, millions of components can be prefabricated in a single IC. Most digital circuits are made of such VLSI (Very Large Scale Integration) IC as its main circuit component with a few decoupling capacitor for supplying clean DC voltage.

It is important to note that any analog signal can be converted to a digital signal (in the form 1s or 0s). Example given below gives a clue about how analog signals can be represented as digital signals,

ANALOG VOLTAGE	DIGITAL VALUE
0 volt	0 0 0 0
1 volt	0 0 0 1
2 volt	0 0 1 0
3 volt	0 0 1 1
4 volt	0 1 0 0
5 volt	0 1 0 1
6 volt	0 1 1 0
7 volt	0 1 1 1
8 volt	1 0 0 0
9 volt	1 0 0 1
10 volt	1 0 1 0

Details of how this conversion is done is discussed in further lessons.

Digital systems offer the following advantages over analog systems

- Easier to design
- Information storage is easy
- Accuracy and precision are greater
- Programmable
- Circuitry can be fabricated on IC chips more easily
- High speed functions

The operations carried out using digital signals are called Logic operations. Example of Logic operation are AND, OR & NOT

- the circuit output should be Logic-1 if atleast any one of the two inputs is Logic-1.

A circuit that performs such a logical operation is

called as a **OR** gate.

- the circuit output should be Logic-1 only when both the inputs are Logic 1's.

A circuit that performs such a logical operation is called as a **AND** gate.

- the circuit output should be inverse of the input. If the input is Logic-1, then the output should be Logic-0 and vice-versa.

A circuit that performs such a logical operation is called as a **NOT** gate.

Every logic operation, even the most extensive and the most complicated - can be reduced to combinations of the above said three basic logic functions. By combining these three operations, several other functions such NAND, NOR and so on (discussed in further paragraphs).

These basic functional circuits are called Gates, such as OR gate, AND gate and NOT gate. The practical implementation of logic operations is effected by logic circuits. In the meantime, a large number of circuit families have been produced in integrated circuit technology. The starting point of standard development was the TTL (Transistor-Transistor-Logic) family (earlier to it was the RTL and DTL families), from which several other families with improved properties have been derived. The TTL family of gates have defined voltage levels and permissible tolerances. Some of the important terminologies associated with digital ICs are given below;

TERMINOLOGY OF DIGITAL IC'S

Saturated logic gate

A form of logic gate in which one output state is the saturation voltage of a transistor.

Example

Resistor Transistor Logic (RTL), Diode Transistor Logic (DTL) and Transistor Transistor Logic (TTL).

Unsaturated logic or current mode logic gate

A form of logic with transistors outside the saturated region.

Example

Current Mode Logic (CML) and Emitter Coupled Logic. This has ultra-fast switching speed and low logic swing.

Operating voltages

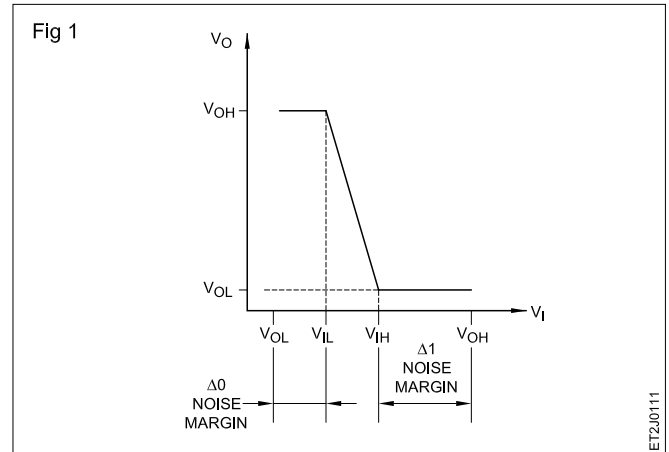
The various operating voltages of a logic gate can be understood with the help of the transfer characteristic of the gate as shown in Fig 1.

VOH - The minimum voltage which will be available at a gate output when the output is supposed to be at logic '1'.

VIH - The minimum gate input voltage which will unambiguously be accepted by the gate as logic '1'.

VOL - The maximum voltage which will appear at a gate output when the output is supposed to be at logic '0'.

VIL - The maximum gate input voltage which will unambiguously be accepted by the gate as logic '0'.



Logic swing: The difference between the two output voltages ($V_{OH} - V_{OL}$) is known as the logic swing of the circuit. That is,

$$\text{logic swing} = V_{OH} - V_{OL}$$

Noise margin

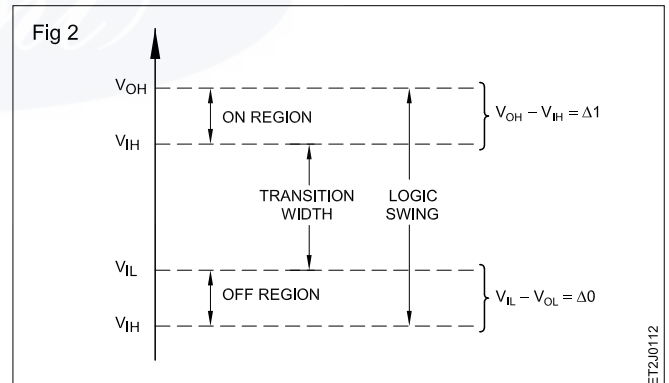
The amount of voltage of extraneous signal which can be tolerated before an output voltage of gate deviates from the allowable logic voltage levels.

The different noise margins of a logic gate can be understood with the help of logic level diagram at Fig 2.

Low-level noise margin

The difference ($V_{IL} - V_{OL}$) is low level noise margin (D_0) $D_0 = V_{IL} - V_{OL}$

High level noise margin



The difference ($V_{OH} - V_{IH}$) is high level noise margin (D_1) $D_1 = V_{OH} - V_{IH}$

Transition width:

From Fig 2, transition width = $V_{IH} - V_{IL}$

It can be seen that an increased noise margin capability is obtained as either V_{OH} and V_{OL} move away from each other or as V_{IH} and V_{IL} move toward each other. With a larger logic swing or a narrower transition width, the noise margins will improve.

Noise immunity

Stray electric and magnetic fields can induce voltages on the connecting wires between logic circuits. These unwanted signals are known as noise and they may cause the voltage at the input to a logic circuit to drop below V_{IH} (min) or rise

above VIL (max), which may lead to unpredictable operation. The noise immunity of a logic circuit refers to the circuit's ability to tolerate noise without causing unwanted changes in the output voltage.

Fan-out

The number of loads connected to a gate is known as fan-out of the gate. The number of load gates need not be a limiting number of load gates need not be a limiting number. It is also known as loading factor. For example, a logic gate that is specified to have a fan-out of 10 can drive 10 standard logic inputs. If this number is exceeded, the output logic-level voltages cannot be guaranteed.

Types of IC package

The ICs come in a wide variety of package types. The factors which determine the type of package are

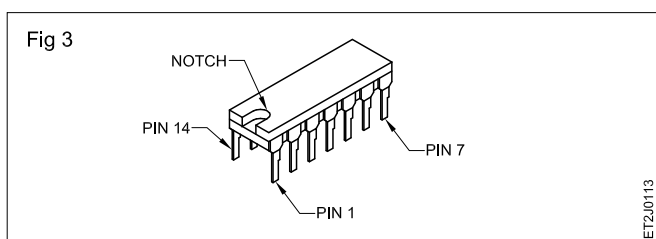
- amount of circuitry contained in the IC
- number of external connections that need to be made to it.
- humidity of the environment, ambient temperature at which the IC is to operate
- method of mounting on the PCB.

DIP [Dual in line package]

The external connecting pins are in parallel rows along the two long edges of the package as shown in Fig 3. In DIP ICs, number of pins varies from 4 to 64 depending on the internal circuitry. For low temperature and low humidity, epoxy plastic packages are used. For high temperature or for devices that dissipate large amount of power, ceramic packages are used.

Ceramic flat package

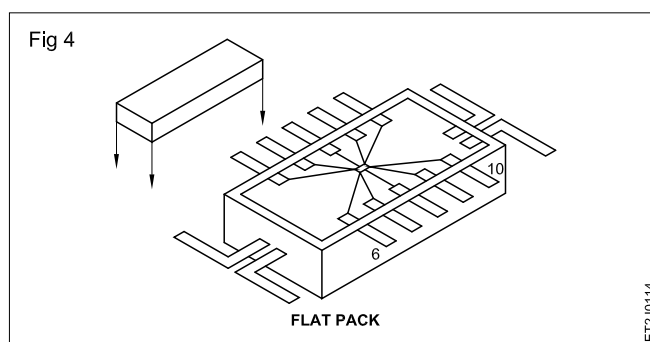
This type of IC packages are hermetically sealed as shown in Fig 4, which means that they are totally immune to the effects of humidity. These packages are often used in



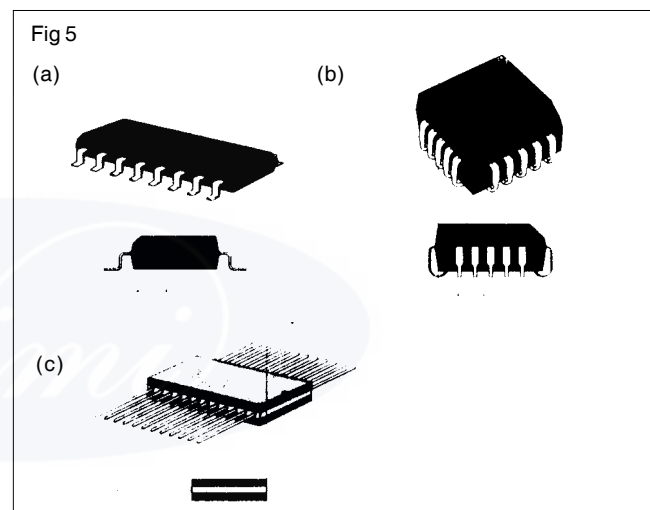
military equipments that they must be able to withstand harsh environments. Pins are counted around the package from notch or dot. These packages are usually mounted in high quality sockets on the circuit board.

Surface mount package

This popular package is similar to the standard DIP except that it is smaller and, as the name implies, its pins are constructed so that it can be soldered directly to metal

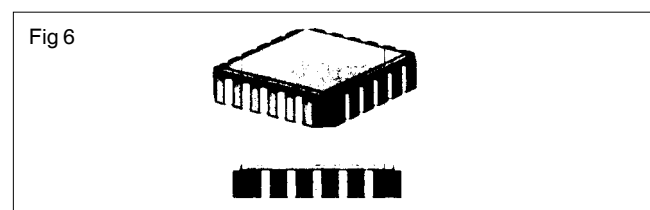


pads on the PCB. One type of package called small out line IC is shown in Fig 5a. Since surface mount packages are soldered on one surface of the circuit board, holes don't have to be drilled on the PCB. Surface mount devices have further advantages, that they are more easily handled by equipment, which automatically mounts components in the correct position on circuit boards during manufacturing. The PLCC (Plastic Leaded chip carrier) type package is shown in Fig 5b. Another variety of SMT package is known as Flat pack is shown in Fig 5c.



Ceramic chip carrier package

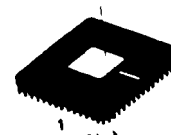
These chips are intended to be clamped into a socket as shown in Fig 6 so that the pads press against contacts which are connected to P.C.B signal lines pin 100 this package is to the right of the notched corner.



Pin grid array package

These ICs are used for VLSI digital circuits such as microprocessor. The number of pins in the array depends on the complexity of the internal circuit. The four corner pin positions are usually left without pins. Common array sizes are 10 x 10, 13 x 13 and 14 x 14, large ICs such as these are put in sockets so they can easily be replaced if the device fails.

Fig 7



S.No.	Complexity	Number of Gates	Application
1	Small-scale integration (SSI)	Fewer than 12	Basic gates
2	Medium-scale integration (MSI)	12 to 99	Flip-flops, register etc.
3	Large-scale integration (LSI)	100 to 9999	Memories, microprocessor
4	Very large-scale integration (VLSI)	10,000 to 99,999	-do-
5	Ultra large-scale integration (ULSI)	100,000 or more	-do-

Logic family

Digital ICs are classified not only by their complexity, logical operation, speed of operation but also by the specific circuit technology to which they belong. The circuit technology is referred to as a digital logic family. Each logic family has its own basic electronic circuit upon which more complex digital circuit and components are developed. The basic circuit in each technology is NAND, NOR or an inverter gate. The electronic components, and material used in the construction of the basic circuit are usually used as the name of the technology. The various logic families of ICs used in electronic circuit are briefly discussed below.

T.T.L Logic family

The word T.T.L is expanded as Transistor-Transistor Logic. In this family ICs are built with transistors. Most standard T.T.L ICs require a power supply voltage between +4.75V and +5.25V to operate properly. The IC(s) of standard T.T.L family are identified by numbers that start with 74 or for military specification devices 54, two or three digits after the 74 or 54 are used to identify the logic functions performed by the device. Some of the 74 series TTL IC numbers with their functions given at Appendix 'D'.

The T.T.L logic family consists of several sub families as shown. The difference between the various T.T.L series are in their electrical characteristics, such as power dissipation, propagation delay and switching speed. They do not differ in the pin assignment or logic operation performed by the internal circuits.

The most popular 7400 series is a line of standard TTL chips. This bipolar family contains variety of compatible SSI and MSI devices. One way to recognise TTL design is the multiple emitter input transistors and the totem pole output transistors. The standard TTL chip has a power dissipation of about 10mw/gate and a propagation delay of around 10ns. The series 74500 is a schottky version having a schottky diode in parallel with collector-base terminals. In this, transistors are prevented from saturating thereby propagation delay is reduced typically to 3ns. By increasing internal resistances and including schottky diodes, low

power schottky diodes numbered from 74LS00 are manufactured limiting the power dissipation to 2mw per gate low power schottky TTL is the most widely used of the TTL types. In this family of devices, a floating input is equivalent to a high input. In electrically noisy environment, floating inputs may pick up enough noise voltage to produce unwanted changes in the output stages and hence inputs should not be kept floating in TTL family. A modified TTL design namely three state TTL allows us to connect outputs directly. Earlier computers used open-collector devices with their bases but the passive pull-up limited the operating speed. These newer devices are much faster and have a control input that can turn off the devices. When this happens the output floats and presents a high impedance to whether it is connected to and hence are widely used for connecting to bases.

E.C.L

Emitter-coupled logic circuits provide the highest speed with propagation delay typically of 5ns. The most common ECL ICs are designated as the 10000 series. E.C.L is used in systems such as super computers and signal processors where high speed is essential. The ECL family IC use is restricted to few applications because of the following reasons.

- The gates in ICs dissipate relatively large amounts of power.
- Needs extra circuitry for gates to operate.
- The -ve power supply voltage and logic levels make ECL gates difficult to interface with other logic family members.

MOS

The Metal Oxide Semiconductor is a unipolar transistor that depends upon the flow of only one type of carrier, which may be either electrons or holes. A p-channel MOS is referred to as PMOS and an N-channel as NMOS. NMOS is the one that is commonly used in circuits with only one type of MOS transistor. MOS technology allows a very large number of circuits to be built in a single IC. It is this technology which has made possible the microprocessors, memories and other LSI devices which are used to build microcomputers.

CMOS

PMOS circuitry or NMOS circuitry can't be used alone for making simple logic gate devices such as done in T.T.L devices, for various reasons. However by building the circuits on the IC by using one PMOS and NMOS transistors connected in complementary fashion, it is possible to produce logic gate devices which have the desired characteristics. That is why these ICs are called complementary metal oxide semiconductors or just CMOS.

CMOS ICs are designated in 4000 series. The family includes logic functions such as those available in T.T.L family. CMOS sub families operate properly with a power supply voltage of +3V to +15V.

By the time the 4000 series of CMOS was developed, most logic designers has become very familiar with the logic functions, part numbers and pin connections of the devices in the standard T.T.L sub family. So as to make CMOS ICs more compatible with T.T.L standards, the CMOS, TTL compatible ICs are made available in 74C00 series. 74HC00 series (high speed) 74HCT00 series, refer Table Performance comparison of CMOS and TTL logic families is given in the Table below:

Technology	CMOS		TTL				
	Silicon Gate	Metal-Gate	Std.	Low-Power Schottky	Schottky	Advanced Low-Power Schottky	Advanced Schottky
Device series	74HC	4000	74	74LS	74S	74ALS	74AS
Power dissipation (mW/gate)							
Static	0.0000025	0.001	10	2	19	1	8.5
AT 100 kHz	0.17	0.1	10	2	19	1	8.5
Propagation delay time (ns) (CL = 15 pF)	8	50	10	10	3	4	1.5
Maximum clock frequency (MHz) (CL = 15 pF)	40	12	35	40	125	70	200
Speed/Power product (pJ) (at 100 kHz)	1.4	11	100	20	57	4	13
Minimum output drive IOL (mA) (VO = 0, .4V)	4	1.6	16	8	20	8	20
Fan-out; LS loads	10	4	40	20	50	20	50
Same-series	*	*	10	20	20	20	40
Maximum input current, IIL (mA) (VI = 0.4V)	±0.001	-0.001	-1.6	-0.4	-2.0	-0.1	-0.5

*Fan-out is frequency dependent

COMPARISON OF TYPICAL NOISE MARGINS:

Noise Margin	HCMOS (V)	STD TTL (V)	LS TTL (V)	S TTL (V)	AS TTL (V)
VNH	1.4	0.4	0.7	0.7	0.7
VNL	0.9	0.4	0.4	0.4	0.4

Digital I.C numbering system

Number and letters on IC packages identify the logic family and the logic function of a device. In addition to these numbers and letters, an IC may have numbers and letter which indicate manufacturers name, the factory where the device was manufactured, the year and month the device manufactured, the package type and a code which indicates how thoroughly the device was tested.

Ex: 74HCT00N

74 HCT 00 N

XXX XXX X

Letter codes for common package type

N = Plastic dip

J = Ceramic dip

D = Glass/metal dip

W = Flat pack

Manufacturer code

Code	Manufacturer
AM	Advanced microdevice
CD	GE/RCA
DM	National semiconductor
F	Fair child
GD	Gold star
H	Harris
HD	Hitachi
IM	Intersil
KS	Samsung
LR	Sharp
M	SGS
MC	Motorola
MM	Monolithic memories
MN	Panasonic
N	Signetics
P	Intel
SN	Texas instruments
SP	SPI
US	Sprague
TC	Toshiba

Component Handling Precautions

CMOS ICs need careful handling. Obviously one should take reasonable care in handling all components, especially nowadays when so many are of small size, but certain components can be damaged by high voltage static charges. It is easy for these to occur without evidence of their presence, because they are generated by friction between insulating materials, and because so many different plastic materials with very low conductivity are in common everyday use. This voltage which is dangerous to electronic devices.

CMOS devices can be damaged by voltages as low as 250V, so the first precaution to observe is to ensure that the handler is not carrying a static charge. An operative handling such devices in the course of their daily work may have an earthed mat on the bench and an earthed wrist strap to allow any charge to leak away. It is however wise to earth one's body momentarily before handling a CMOS IC and this can be done by touching the metal casing of any earthed electrical appliance which is plugged in but not switched on e.g. a cold soldering iron.

CMOS ICs are supplied in some form of conductive packing, which may be simply aluminium foil, or black (carbon filled) plastic foam. Shielding bags are sometimes used, but the plastic tubes marked ANTISTATIC are not very effective and ICs in these should be handled with extra care. A sensible way to ensure that devices are not damaged is to use IC holders and to insert the devices only when construction is complete and the layout has been checked. Plastic insertion and removal tools are not expensive and are preferable to fingers for manipulating the devices. No device should be inserted while power is applied to the Circuits and obviously one should make sure that the device is inserted the right way round.

Here provided a list of components and the voltages which can damage them, and also some details about protecting such devices:

Device Type	Minimum damaging voltage
VMOS	30
MOSFET	100
EPROM	100
JFET	140
Op-amp	190
CMOS	250
Schottky diodes	300
Film resistors	300
Bipolar transistors	380
SCRs	680
Schottky TTL	1000

Most CMOS ICs are equipped with internal protective diodes which limit voltages at the inputs to a safe level, but care is still advisable in handling them. If they have to be soldered in directly, the VDD pin should be connected before any earth connection is made.

Some beginners are worried about the possibility of component damage during soldering operations. It is true that some components are affected by heat, for instance, resistors change their value when heated, and extreme heat can damage Semiconductors such as diodes and

transistors. The best way to avoid such damage is to acquire sound soldering technique. If the surfaces being soldered are clean, the soldering iron need only be applied to the joint for a short time, and the temperature rise in the body of the component will be too small to cause any damage. One can buy soldering stunts to clip on transistor or other leads during soldering operations, but they are really unnecessary - a pair of flat nosed pliers can be applied where one needs to avoid heat being conducted to a sensitive device.



Number systems and conversions.

Objectives: At the end of this lesson you shall be able to

- differentiate between different number systems like decimal, octal, binary and hexadecimal and conversion between them and different types of codes
- explain NOT gate using transistor.

INTRODUCTION

When we hear the word 'number' immediately we recall the decimal digits 0,1,2....9 and their combinations. Modern computers do not process decimal numbers. Instead, they work with binary numbers which use the digits '0' and '1' only. The binary number system and digital codes are fundamental to digital electronics. But people do not like working with binary numbers because they are very long when representing larger decimal quantities. Therefore digital codes like octal, hexadecimal and binary coded decimal are widely used to compress long strings of binary numbers.

Binary number systems consists of 1s and 0s. Hence this number system is well suited for adopting it to the digital electronics.

The decimal number system is the most commonly used number system in the world. It uses 10 different characters to show the values of numbers. Because this number system uses 10 different characters it is called base-10 system. The base of a number system tells you how many different characters are used. The mathematical term for the base of a number system is radix.

The 10 characters used in the decimal number systems are 0,1,2,3,4,5,6,7,8,9.

Positional notation and weightage

A decimal integer value can be expressed in units, tens, hundreds, thousands and so on. For example decimal number 1967 can be written as $1967 = 1000 + 900 + 60 + 7$. In powers of 10, this becomes

				$1 \times 10^3 = 1000$
10^3	10^2	10^1	10^0	$9 \times 10^2 = 900$
				$6 \times 10^1 = 60$
1	9	6	7	$7 \times 10^0 = 7$
				1967

i.e. $[1967]_{10} = 1(10^3) + 9(10^2) + 6(10^1) + 7(10^0)$

This decimal number system is an example of positional notation. Each digit position has a weightage. The positional weightage for each digit varies in the sequence $10^0, 10^1, 10^2, 10^3$ etc starting from the least significant digit.

The sum of the digits multiplied by their weightage gives the total amount being represented as shown above.

In a similar way, binary number can be written in terms of weightage.

To get the decimal equivalent, then the positional weightage should be written as follows.

$$[1010]_2 = 1(2^3) + 0(2^2) + 1(2^1) + 0(2^0)$$

$$= 8 + 0 + 2 + 0$$

$$[1010]_2 = [10]_{10}$$

Any binary number can be converted into decimal number by the above said positional weightage method.

Decimal to Binary conversion

Divide the given decimal number by 2 as shown below and note down the remainder till you get the quotient - zero.

Example

	0	
2	1	1 → MSB
2	2	0
2	4	0
2	8	0
2	17	1
2	34	0 → LSB

The remainder generated by each division form the binary number. The first remainder becomes the LSB and the last remainder becomes the MSB of binary number.

Therefore, $[34]_{10} = [100010]_2$

Counting binary number

To understand how to count with binary numbers, let us see how an odometer (Km indicator of a car) counts with decimal numbers,

The odometer of a new car starts with the reading 0000.

After traveling 1km, reading becomes 0001.

Successive km produces 0002, 0003 and so on upto 0009

At the end of 10th km, the units wheel turns back from 9 to 0, a tab on this wheel forces the tens wheel to advance by 1. That is why the number changed from 0009 to 0010. That is, the units wheel is reset to 0 and sent a carry to the tens wheel. Let us call this familiar action as reset and carry. The other wheels of odometer also reset and carry. For instance, after covering 999km, the odometer shows 0999.

After the next km, the unit wheel resets and carries, the tens wheel resets and carries, the hundreds wheel resets and carries and the thousands wheel advances by 1 to get the reading 01000.

Binary odometer

Visualize a binary odometer, a device whose wheels have only two digits 0 and 1. When each wheel turns, it displays 0 then 1 and then back to 0 and the cycle repeats. A four digit binary odometer starts with 0000.

After 1km, it indicates - 0001.

The next km forces the units wheel to reset and sends carry. So the number changes to 0010.

The third km results in 0011.

After 4km, the units wheel resets and sends carry, the second wheel resets and sends carry and the third wheel advances by 1. Hence it indicates 0100.

Table below shows all the binary numbers from 0000 to 1111 equivalent to decimal 0 to 15.

Decimal	Binary
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111

Addition of binary numbers

Sum	Carry
0 + 0 = 0	0
1 + 0 = 1	0
0 + 1 = 1	0
1 + 1 = 0	1 (one plus one is equal to zero with carry one)

Ex: 1

$$\begin{array}{r} 10 \\ + 11 \\ \hline 101 \end{array}$$

Ex: 2

$$\begin{array}{r} 1 + 1 + 1 = 1 \quad (\text{one plus one plus one is equal to one with carry one}) \\ + 1 \\ \hline 10 \\ + 1 \\ \hline 11 \end{array}$$

Hexadecimal number system: In hexadecimal system there are 16 characters. They are 0,1,2,3,4,5,6,7,8,9, A,B,C,D,E,F where A=10, B=11, C=12, D=13, E=14, F=15 in decimal. In this system, the base is 16. This system is mainly used to develop programmes for computers.

For Example

$$[23]_{16} = [35]_{10}; 16^1 \times 2 + 16^0 \times 3 = 32 + 3 = 35;$$

$$[2C]_{16} = [44]_{10}; 16^1 \times 2 + 16^0 \times 12 = 32 + 12 = 44;$$

Decimal to hexadecimal conversions

The conversion of decimal to hexadecimal is similar to binary conversion. Only difference is that divide the decimal number successively by 16, and note down the remainder.

$$\begin{array}{r} 0 \\ 16 \overline{) 1} \quad 1 \longrightarrow \text{MSB} \\ 16 \overline{) 27} \quad 11 \text{ or B} \\ 16 \overline{) 432} \quad 0 \longrightarrow \text{LSB} \end{array}$$

$$[432]_{10} = [1B0]_{16}$$

Hexadecimal to Decimal

This conversion can be done by putting it into the positional notation.

$$\begin{aligned} \text{Ex: } 223A_{16} &= 2 \times 16^3 + 2 \times 16^2 + 3 \times 16^1 + A \times 16^0 \\ &= 2 \times 4096 + 2 \times 256 + 3 \times 16 + 10 \times 1 \\ &= 8192 + 512 + 48 + 10 \\ &= 8762_{10} \end{aligned}$$

OCTAL NUMBER

The octal number system provides a convenient way to express binary numbers. It is used less frequently compared to hexadecimal in conjunction with computers and microprocessors to express binary quantities for input and output purposes.

The octal number system is compared of digit symbols such as right symbols such as 0,1,2,3,4,5,6,7.

Since there are 8-symbols, radix or base is 8. Positional weightage is $8^3, 8^2, 8^1, 8^0$.

To distinguish octal numbers from other number systems subscript 8 is used as follows:

Ex: $(15)_8 \sim (13)_{10}$

Octal Decimal

Octal to Decimal conversion

	0		
8	1	1	→ MSB
8	13	5	→ LSB

$$(13)_{10} = (15)_8$$

As in other number systems, each digit should be multiplied by its positional weightage and added to get decimal equivalent.

Convert $(2374)_8$ into decimal number

Positional weightage : $8^3, 8^2, 8^1, 8^0$

Octal number 2 3 7 4

$$(2374)_8 = (2 \times 8^3) + (3 \times 8^2) + (7 \times 8^1) + (4 \times 8^0)$$

$$= (2 \times 512) + (3 \times 64) + (7 \times 8) + (4 \times 1)$$

$$= 1024 + 192 + 56 + 4$$

$$(2374)_8 = (1276)_{10}$$

Decimal to octal conversion

A method of converting a decimal number to an octal number is the repeated division by 8, each successive division by 8 yields a remainder that becomes a digit in the equivalent octal number. The first remainder generated is the least significant digit (LSD).

$$(359)_{10} = (547)_8$$

	0		
8	5	5	→ MSB
8	44	4	
8	359	7	→ LSB

Octal to binary

Each octal digit can be represented by a 3-bit binary number, because of this it is very easy to convert from octal to binary. Each octal digit is represented by three bits as shown in the table.

Octal	0	1	2	3	4	5	6	7
digit								
Binary	000	001	010	011	100	101	110	111

To convert each octal number to a binary, simply replace each octal digits with the corresponding binary bits.

Example

$$1 \quad (25)_8 = (\quad)_2$$

$$\begin{array}{cc} 2 & 5 \\ 010 & 101 \end{array}$$

$$(25)_8 = (010101)_2$$

$$2 \quad (7526)_8 = (\quad)_2$$

$$\begin{array}{cccc} 7 & 5 & 2 & 6 \\ 111 & 101 & 010 & 110 \end{array}$$

$$(7526)_8 = (111101010110)_2$$

Binary to octal

Conversion of a binary number to an octal number is the reverse of the octal-to-binary conversion. The procedure is as follows.

- 1 Start with the right most group of three bits and moving from right to left, convert each 3-bit group to the equivalent octal digit.
- 2 If there are not three bits available for the left most group, add either one or two zero's to make complete group. These leading zero's will not affect the value of the binary number.

Example

$$(110101)_2 = (\quad)_8$$

$$\begin{array}{cc} 110 & 101 \\ 6 & 5 \end{array} = (65)_8$$

$$(11010000100)_2 = (\quad)_8$$

$$\begin{array}{cccc} 011 & 010 & 000 & 100 \\ 3 & 2 & 0 & 4 \end{array} = (3204)_8$$

BCD (Binary Coded Decimal)

Binary Coded Decimal (BCD) is a way to express each of the decimal digits with a binary code, since there are only ten code groups in the BCD system, it is very easy to convert between decimal and BCD. Because decimal system is used for read and write, BCD code provides an excellent interface to binary systems. Examples of such interfaces are keypad inputs and digital readouts.

8421 code

The 8421 code is a type of binary coded decimal (BCD), binary coded decimal means that each decimal digit, 0 through 9 is represented by a binary code of four bits. The designation 8421 indicates the binary weights of the four

bits ($2^3, 2^2, 2^1, 2^0$). The ease of conversion between 8421 code numbers and the familiar decimal numbers in the main advantage of this code. All you have to remember are the ten binary combinations that represents the ten decimal digits as shown in Table.

Decimal digit	0	1	2	3	4	5	6	7	8	9
BCD	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001

The 8421 code is the pre-dominant BCD code, and when we refer to BCD, we always mean the 8421 code unless otherwise stated.

Invalid code

You should realize that with four bits, sixteen numbers (0000 through 1111) can be represented, but in the 8421 code only ten of these are used. The six code combinations that are not used 1010, 1011, 1100, 1101, 1110 and 1111 are invalid in the 8421 BCD code.

To express any decimal number in BCD, simply replace each decimal digit with the approximate 4-bit binary code.

Example

1 $(35)_{10} = (?)$ 8421 code

3 5
0011 0101 = 00110101

2 $(2458)_{10} = (?)$ 8421 code

2 4 5 8
0010 0100 0101 1000 = 0010010001011000

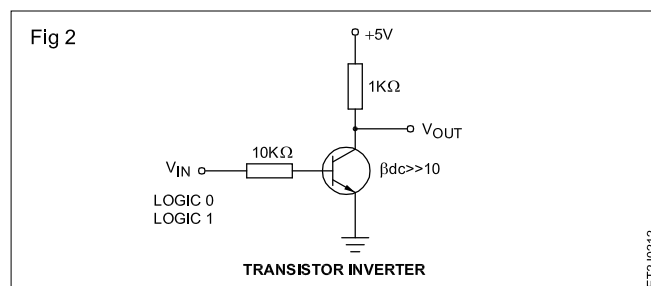
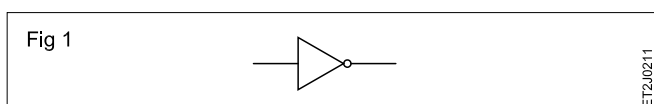
There are many specialized codes used in digital system other than BCD code. Some codes are strictly numeric, like BCD and others are alphanumeric which are used to represent numbers, letters, symbols and instructions.

The commonly used codes other than BCD codes are

- 1 Gray code
- 2 Excess 3 code
- 3 ASCII code - American, Standard code for Information interchange
- 4 Alphanumeric code

Inverters (NOT Gate)

An inverter is a gate with only one input signal and one output signal. The output state is always the opposite of the input state. Logic symbol is shown in Fig 1.



Transistor inverter

The above circuit shows the transistor inverter circuit. The circuit is a common emitter amplifier which works in saturation or in cut off region depending upon the input voltage. When V_{in} is in low level, say less than the transistor cut in voltage 0.6V in silicon type, the transistor goes to cut off condition and the collector current is zero. Therefore, $V_{out} = +5V$ which is taken as high logic level. On the other hand, when V_{in} is in high level, the transistor saturates and $V_{out} = V_{sat} = 0.3V$ i.e low level.

The table summarizes the operation

V_{in}	V_{out}
Low(0)	High(1)
High(1)	Low(0)

The logic expression for the inverter is as follows: If the input variable is 'A' and the output variable is called Y, then the output $Y = A$.

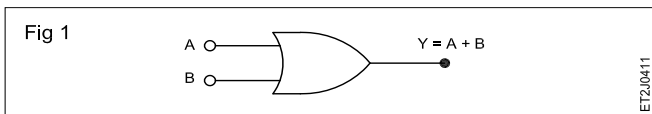
OR, NOR gate circuits and RS flip flop

Objectives: At the end of this lesson you shall be able to

- explain OR gate using diode and its truth table
- explain T.T.L OR gate IC 7432
- explain NOR gate and its truth table.

OR GATE

The output of an OR will be in 1 state if one or more of the inputs is in 1 state. Only when all the inputs are in 0-state, the output will go to 0-state. Fig 1 shows the schematic Symbol of an OR Gate :



The boolean expression for OR gate is $Y = A + B$.

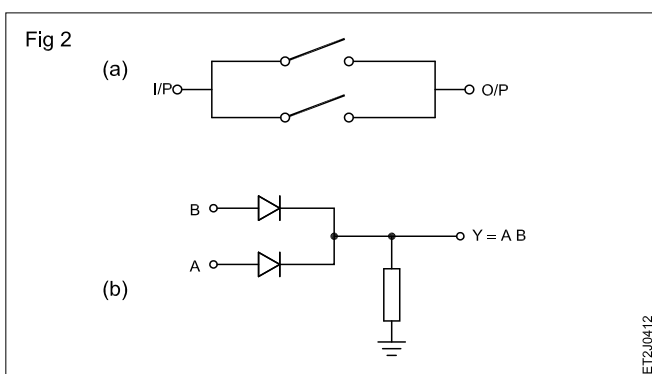
The equation is to be read as Y equals A ORed B. Two-input truth table given below is equivalent to the definition of the OR operation.

Truth table for OR gate

A	B	$Y = A + B$
0	0	0
1	0	1
0	1	1
1	1	1

Electrical equivalent circuit

The Fig 2a shows the electrical equivalent circuit of an OR gate. It is evident that if any one of the switch is closed, there will be output.



2 in-input OR gate using diode

The Fig 2b shows one way to build a 2-input OR gate, using diodes. The inputs are labeled as A and B, while the output is Y.

Assume logic 0 = 0V (low)
 logic 1 = +5V (high)

Since this is a 2 input OR gate, there are only four possible cases,

Case 1: A is low and B is low. With both the input voltage low, both the diodes are not conducting. Therefore the output Y is in low level.

Case 2: A is low and B is high, The high B input voltage (+5V) forward biases the lower diode, producing an output voltage that is ideally +5V (actually +4.3V taking the diode voltage drop 0.7V into consideration). That is, the output is in high level. During this condition, the diode connected to input A is under reverse bias or OFF condition.

Case 3: A is high and B is low, the condition is similar to case 2. Input A diode is ON and Input B diode is OFF and Y is in high level.

Case 4: A is high, B is high. With both the inputs at +5V, both diodes are forward biased, since the input voltages are in parallel, the output voltage is +5V ideally [+4.3V to a second approximation]. That is, the output Y-is in high level.

OR gates are available in the IC form. IC7432 is a T.T.L OR gate IC having 4 OR gates inside it. For pin diagram refer Appendix 'D'.

Simple application of OR gate

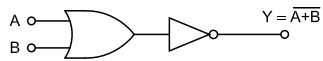
Intrusion detection

Simplified portion of an intrusion detection and alarm system is two windows and a door. The sensors are magnetic switches that produce a high(1) output when windows and doors are opened and a low(0) output when closed. As long as the windows and the door are secured, the switches are closed and all three of the OR gate inputs are in low(0). When one of the windows or the door is opened, a high(1) output is produced on that input of the OR gate and the gate output goes high. It then activates an alarm circuit to warn of the intrusion.

NOR Gate

In the Fig 3 the output y of the circuit equals the complement of A OR B, because the circuit is an OR gate followed by a NOT gate. To obtain high output [Logic-1], both the inputs should be tied to low input [Logic-0]. For the rest of the other three possibilities, output will be zero, the combination of this OR and NOT gate is called as NOR gate.

Fig 3



ET2J0413

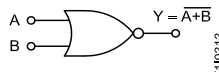
We can define a NOR gate as follows:

The output of a NOR gate is 0, even if one of the inputs is in logic-1. Only when both the inputs are in logic-0, the output is in logic-1.

Truth table

A	B	A+B
0	0	1
0	1	0
1	0	0
1	1	0

Symbol:



AND, NAND and XOR gate circuits

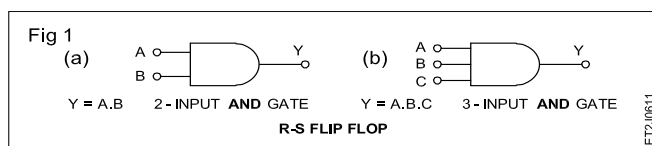
Objectives: At the end of this lesson you shall be able to

- explain AND gate using diodes and its truth table
- explain NAND gate and its truth table
- explain AND gate application
- explain logic pulser using NAND gates.

AND gates

The AND gate has two or more inputs but only one output. All input signals must be held high to get a high output. Even if one of the inputs is low, the output becomes low.

AND gate symbols for 2 input and 3 input gates are shown in Fig 1a and 1b.



Truth table

Two input AND gate

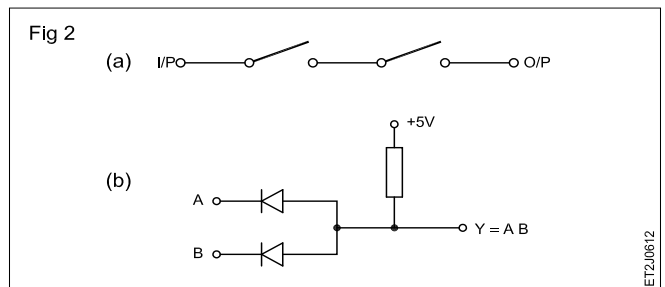
A	B	Y=AB
0	0	0
0	1	0
1	0	0
1	1	1

Three input AND GATE

A	B	C	Y=ABC
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Electrical equivalent circuit of an AND gate

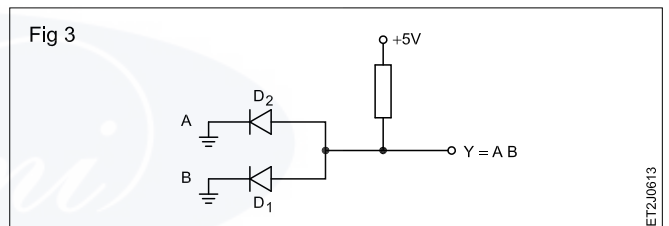
The output is available only when both the switches are closed. IC7408 is a T.T.L quad AND gate IC. (Refer data book for pin diagram). The electrical equivalent of AND gate and AND gate using diodes are shown in Fig 2a and 2b.



Two input AND gate using diode

I condition

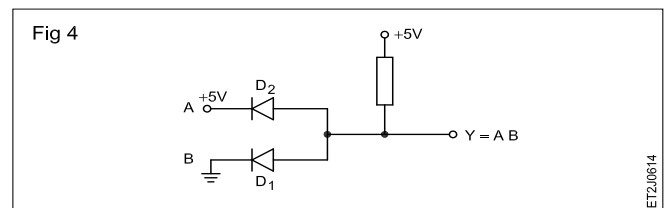
A=0, B=0, Y=0 as shown in Fig 3.



During the above condition I/P A and B are connected to ground to make logic low inputs. During this condition, both the diodes conduct, and pulls the O/P Y to logic-0.

II condition

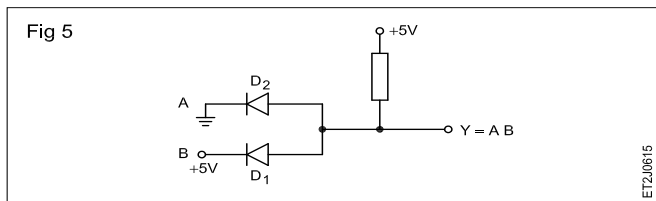
A=0, B=1, Y=0 as shown in Fig 4.



In the II condition shown in the figure above, diode D_1 is connected logic-0 input and diode D_2 is connected to +5V [Logic high]. Diode D_1 is in forward bias and conducts. Diode D_2 is having equal potential (+5V) at anode and cathode. So potential difference between anode and cathode is 0. Hence diode D_2 does not conduct. The output Y is pulled down to logic zero, since D_1 is conducting.

III condition

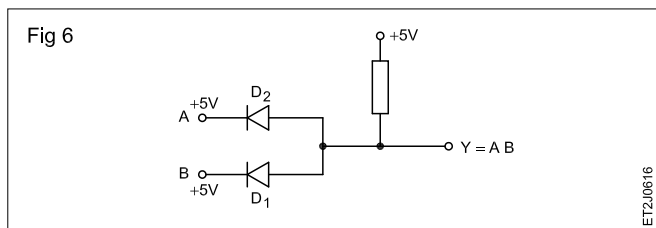
A=1, B=0, Y=0 as shown in Fig 5.



The III condition is similar to the II condition. D_2 is forward biased. D_1 is reverse biased. Hence output Y is pulled to logic-0.

IV condition

A=1, B=1, Y=1 as shown in Fig 6.



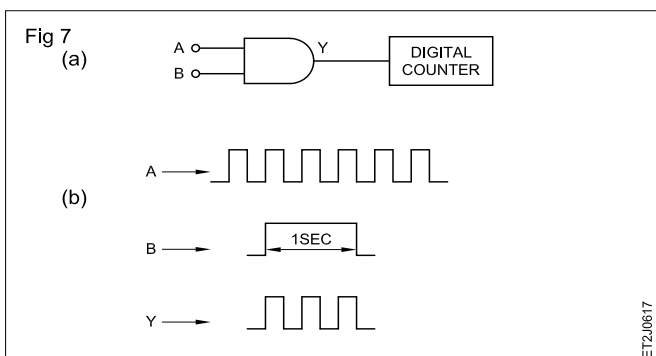
In this condition both the diodes are reverse biased. So both the diodes act as open circuit. Therefore output y is +5V i.e y is in logic-1 condition.

AND gate as an Enable/Inhibit device

A common application of the AND gate is to enable (i.e to allow) the passage of a signal (pulse waveform) from one point to another at certain times and to inhibit (prevent) the passage at other times.

In Fig 7a AND gate controls the passage of a signal (waveform A) to a digital counter. The purpose of this circuit is to measure the frequency of waveform 'A'. The enable pulse has a width of precisely 1 second. When the enable pulse applied at B is high, waveform A passes through the gate to the counter, and when the enabled pulse is low, the signal is prevented (inhibited) from passing through. Refer Fig 7b for the waveforms of the above process.

During the 1 second interval of the enabled pulse, a certain number of pulses in waveform A pass through the AND gate to the counter. The number of pulses counted by the counter is equal to the frequency of the waveform A. For example, if 1000 pulses pass through the gate in the 1 second interval of the enabled pulse, there are 1000 pulses/sec. That is, frequency is 1000Hz.



NAND gate

An AND gate followed by a NOT gate forms the NAND gate as shown in Fig 8a. In this gate to get a low output (logic=0), all the inputs must be in high state and to get high output state, any one of the inputs or both inputs must be in low state.

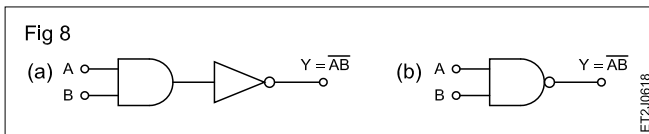


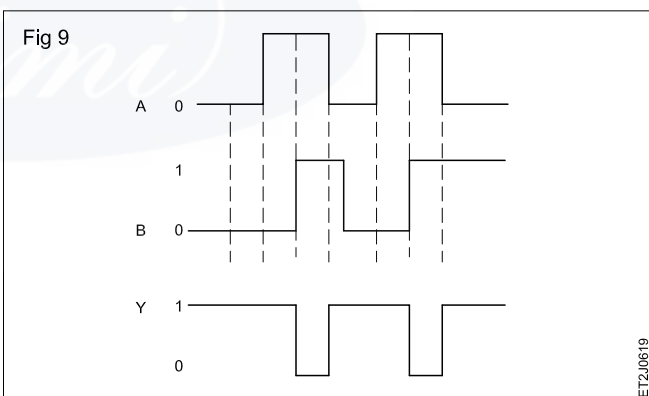
Fig 8b is the standard symbol for a NAND gate. The inverter triangle has been deleted and the bubble is moved to the AND-gate output.

Truth table for NAND gate

A	B	$Y = \overline{AB}$
0	0	1
0	1	1
1	0	1
1	1	0

Pulsed operation

Output waveform Y is low only for the time intervals when both inputs A and B are high as shown in the timing diagram Fig 9.

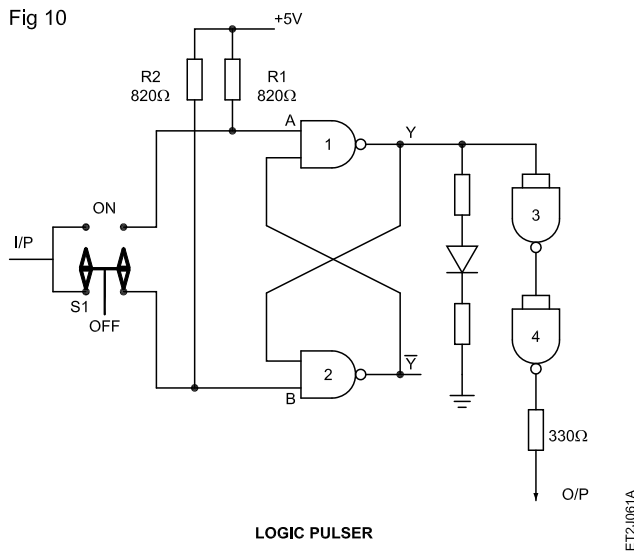


Logic pulser

The Fig 10 shows the circuit diagram of logic pulser, the circuit essentially consists of NAND gates connected dibouncer circuit and its output is Double inverted. The LED indicates, pulses ON or OFF status.

When switch S1 is not pressed, (OFF position) B input of NAND gate No.2 is grounded, hence its output $\overline{Y}$ is forced to go logic HIGH. This HIGH output is feedback to NAND gate 1, A input of NAND gate 1 is also held HIGH through R_1 resistor (820W) and thus the output of NAND gate-1 'Y' is at low. This logic low output keeps LED in OFF condition and this logic low is again double inverted at the logic pulser tip through NAND gate 3 and 4 to get logic low level at pulser tip.

Fig 10



When S1 is pressed to ON, A input of NAND gate is forced to go logic-low. Hence the output of this NAND gate is forced to go logic-HIGH. Therefore the 'Y' output is at logic-1, so LED glows and a logic-HIGH appears at probe tip. Also note that with HIGH at Y output, the inputs of NAND gate 2 are also at logic-HIGH and the output of NAND gate-2 is forced to go low. As long as switch S1 is at ON position the probe tip is HIGH. When it is released it springs back to OFF position, and the output returns to a logic-LOW condition.



R.S Flip-flops and D Flip-flop using NAND gate

Objectives: At the end of this lesson you shall be able to

- logic diagram for the given Boolean equations
- simplify the logic diagram using Boolean algebra
- explain D flip-flop and its truth table
- clocked D flip-flop and its truth table
- explain the difference between edge triggering and level triggering and types of edge trigger
- explain wave shaping circuits used for clocking
- why do you need to know about karnaugh maps
- how karnaugh map is used for boolean equation.

A flip-flop is a digital circuit that has two stable states. It remain in one of these states until triggered into the other.

Flip-flops are used to store binary information. Digital memory circuits that can store bits of data are an essential part of any computer system.

RS flip flops

The most basic type of flip flop is the reset/set type, hence it is known as RS flip flop.

The basic RS flip-flop can be constructed from either two NOR gates or two NAND gates. The circuit symbols is shown in Fig 1. Fig 1a shows RSF/F with active HIGH inputs.

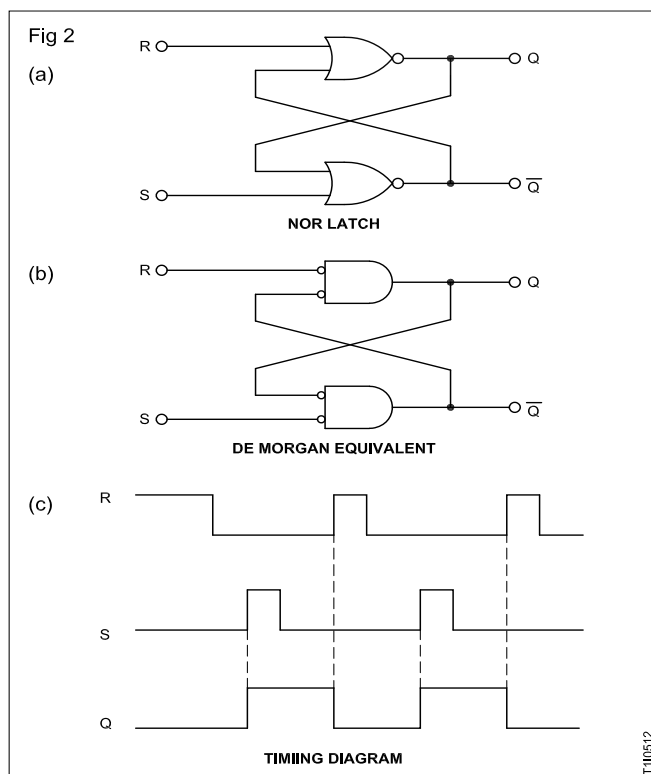
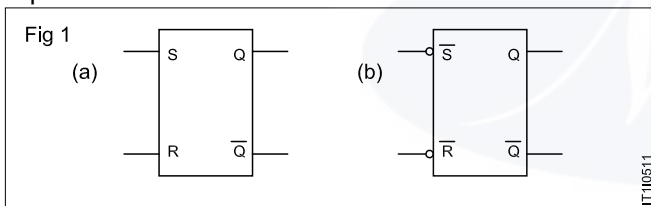
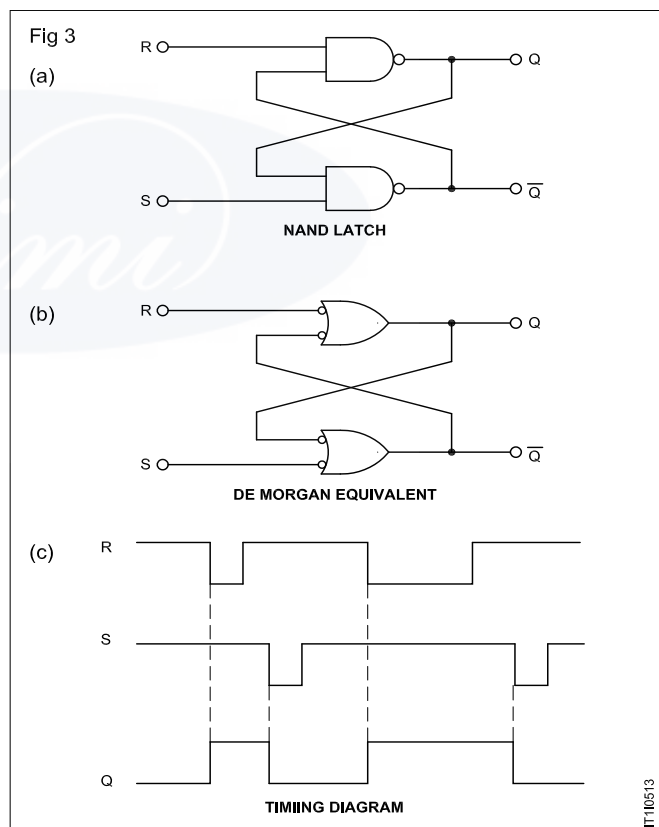


Fig 1b shows RSF/F with active LOW inputs. The NOR gate latch and NAND gate latch both are shown in Fig 2 and Fig 3 respectively.

NOR latch

From Fig 2, the two NOR gates are cross-coupled so that out of one NOR gate is connected to other NOR gate input and vice versa.



Truth table for NOR latch

R	S	Q	Comment
0	0	NC	No change
0	1	1	Set
1	0	0	Reset
1	1	*	Race

Truth table for NAND latch

R	S	Q	Comment
0	0	*	Race
0	1	1	Set
1	0	0	Reset
1	1	NC	No change

The NOR latch output are labelled as Q and $\bar{Q}$. The outputs will always be the inverse of each other. From the truth table of NOR latch, it can be summarised as follows.

Condition 1

R=0 S=0, this condition produce the inactive state. Output 'Q' will remain with no change.

Condition 2

R=0 S=1, this condition cause to go to the Q=1 state where it always remain after R returns high. This is known as setting the latch.

Condition 3

R=1 S=0, this condition cause to go to the Q=0 state where the output remain even after S returns HIGH. This is called resetting the latch.

Condition 4

R=1 S=1, this condition produce a race condition. Therefore avoid R=1 and S=1 condition while using a NOR latch.

NAND latch

From the NAND gate latch as shown in Fig 2, The two NAND gates are cross-coupled so that output of one NAND is connected to other NAND gate input and vice versa. The NAND latch outputs are labelled as Q and $\bar{Q}$. These outputs will always be the inverse of each other.

From the truth table, it can be summarized as follows.

Condition 1

R=0, S=0. This condition produce ambiguous results. It should not be used.

Condition 2

R=0, S=1. This condition cause the output to go the Q=1 state where it will remain after R returns high. This is known as setting the latch.

Condition 3

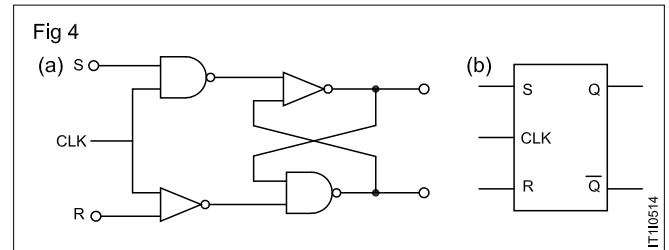
R=1, S=0. This condition cause the output to go the Q=0 state, where the output will remain even after S returns HIGH. This is called clearing or resetting the latch.

Condition 4

R=1, S=1. This condition is the normal resting state and it has no effect on the output state. The Q and $\bar{Q}$ outputs will remain in whatever state they were prior to this input condition.

Clocked RS flip-flop

It is possible to strobe or clock the flip-flop in order to store information (set it or reset it) at any time, and then hold the stored information for any desired period of time. This flip-flop is called a clocked RS flip-flop and is shown in Fig 4a and 4b.



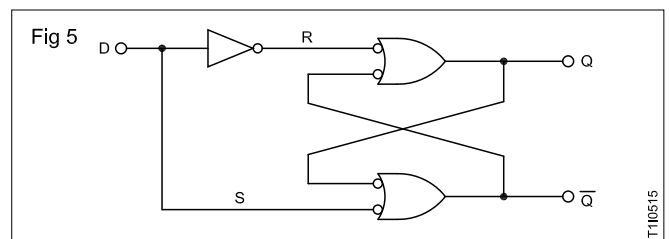
Truth Table

Clock	R	S	Q
0	0	0	NC
0	0	1	NC
0	1	0	NC
0	1	1	NC
1	0	0	NC
1	0	1	1
1	1	0	0
1	1	1	Illegal

For the flip-flop to operate properly there must be a transition from low to high on the clock input, while clock is high, the information on R and S causes the latch to set or reset. Then when clock transitions back to low, this information is retained in the latch. When this high to low transition occurred both R and S inputs were low(0) and thus there was no change of state.

D-flip-flop

The RS flip-flop has two data inputs, R and S. To store a high bit, you need a high S and to store a low bit, you need a high R. Generation of two signals to drive a flip-flop is a disadvantage in many applications. Further more the RS flip-flop is susceptible to a race condition. We will modify the design to eliminate the possibility of a race condition, to overcome the above disadvantage, R.S flip is slightly modified as shown in Fig 5 to have a single input called D-flip-flop.



Unlocked D latch

D	Q
0	0
1	1

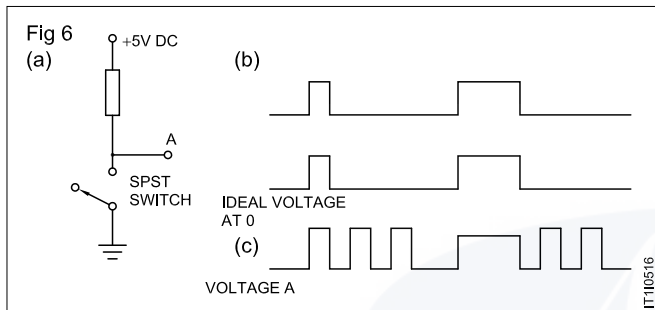
Clock D latch

CLK	D	Q
0	X	NC
1	0	0
1	1	1

Contact bounce circuit

Any mechanical switching device consists of a moving contact arm restrained by some sort of spring system. As a result when the arm is moved from one stable position to the other, the arm bounces as much as hard ball bounces when dropped on a hard surface. The number of bounces that occurs and the period of the bounce differ for each switching device.

In digital system there will be occasions to use mechanical contacts for the purpose of conveying an electrical signal, for example the keyboard of computer. In each case the intent is to apply a high logic level usually [+5V dc] or a low logic level (0 volts DC).



As shown in the Fig 6 above, when the SPST switch is open, the voltage at point 'A' is +5V DC when the switch is closed the voltage at point 'A' is 0 volts.

Ideally the voltage waveform at 'A' should appear as shown in the waveform 6b as the switch is moved from open to closed, or vice-versa.

In actuality, the waveform at point 'A' will appear more or less as shown in waveform (C), as a result of contact bounce. Notice carefully that in this particular instance, even though actual physical contact bounce occurs each time the switch is opened or closed, contact bounce appears in the voltage level at point 'A' only when the switch is closed.

If the voltage at point 'A' is applied to the input of a T.T.L circuit, the circuit will respond properly when the switch is opened, since no contact bounce occurs. However when the switch is closed, the circuit will respond as if multiple signals were applied, rather than the single-switch closer-intended.

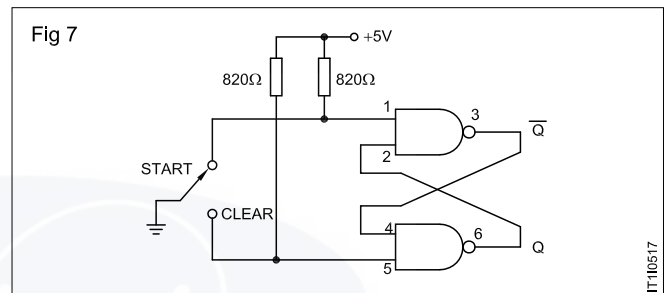
R.S.Latches are often used as switch debouncers. Whenever a switch is thrown from open to the closed position.

Bounce less switch

Floating T.T.L inputs are equivalent to high inputs. With the switch in the start position, Pin 1 is low and pin 5 is high therefore (pin no.3) is high and Q (pin no.6) is low. When the switch is thrown to the clear position, pin no.1 goes high, as shown in Fig 7. Because of contact bounce, pin 5 goes alternately low and high for a few milliseconds before sets in the low state. The first time pin 5 goes low, the latch sets, Q going high and $\bar{Q}$ going low. Subsequent bounces have no effect on Q and $\bar{Q}$ because the latch stays set.

When the switch is thrown back to start, pin 1 bounces low and high for a while. The first time pin 1 goes low, Q goes to low and $\bar{Q}$ going high, later bounces have no effect on Q and $\bar{Q}$.

The Fig 8 shows one way to build a D, latch because of the inverter databit 'D' drives the 'S' input of a NAND latch and the complement of D, drives the 'R' input. Therefore a high 'D' sets the flip-flop, low 'D' resets the flip-flop. Most important thing is no race condition.

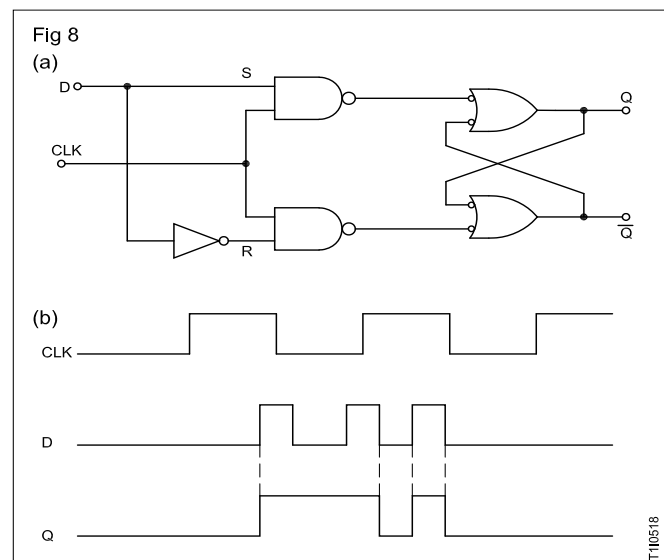


Truth table for D-latch

D	Q	$\bar{Q}$
1	1	0
0	0	1

Clocked D-flip-flop

The Fig 8a shows the level clocked D type flip-flop. A low clock disables the input gates and prevents the latch from changing states, in other words, while clock is low, the latch is in the inactive state D controls the output, A high D sets the latch, while a low D resets it.



Truth table for level clocked D flip flop

Clk	D	Q
0	X	NC
1	0	0
1	1	1

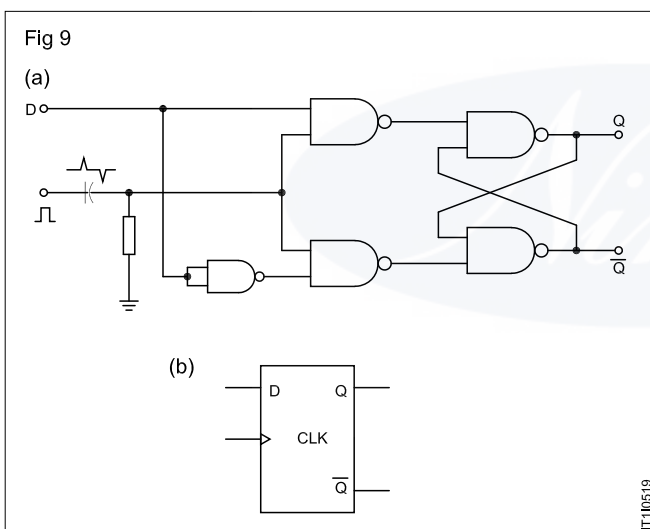
The truth table summarizes the operation 'X' represents a don't care condition, it stands for either 0 or 1, while clock is low the output can't change, no matter what 'D' is, when clock is high, the output equals the input. $Q = D$.

Edge triggering versus level clocking

When a circuit is edge triggered, the output can change only on the rising or falling edge of the clock. Edge triggered D - F/F using discrete gate is shown in Fig 9a and the circuit symbol is shown in Fig 9b.

When the circuit is level clocked, the output can change while the clock is high or low.

With the edge triggering, the output can change only at one instant during the clock cycle. With level clocking, the output can change during the entire period the level of the clock is maintained.



Boolean algebra

Boolean algebra is a convenient and systematic way of expressing and analysing the operation of logic circuits.

Truth table

Edge triggered D-Flip-flop

CLK	D	Q
0	X	NC
1	X	NC
-	X	NC
-	0	0
-	1	1

Variable

A variable is a symbol (usually an italic uppercase letter) used to represent a logical quantity. Any single variable can have a 1 or 0 value.

Ex: A,B,C,D or X,Y,Z etc

Complements

The complement is the inverse of a variable and is indicated by a bar over the variable.

Ex: The complement of A is $\bar{A}$, the complement of A is read as "A bar".

Literal

A literal is a variable or the complement of a variable.

Boolean addition

$$0 + 0 = 0$$

$$0 + 1 = 1$$

$$1 + 0 = 1$$

$$1 + 1 = 1$$

In Boolean algebra, a sum term is a sum of literals. In logic circuits, a sum term is produced by an OR operation with NO AND operation involved.

Ex: $A+B$, $A+\bar{B}$, $\bar{A}+B$

A sum term is equal to 1 when one or more of the literals in the term are 1. A sum term is equal to 0 if and only if each of the literal is 0.

Boolean multiplication

Boolean multiplication is equivalent to the AND operation and the basic rules are as follows.

$$0.0 = 0$$

$$1.0 = 0$$

$$0.1 = 0$$

$$1.1 = 1$$

In Boolean algebra a product term is the product of literals. In logic circuits a product term is produced by an AND operation with NO OR operations involved.

Ex: AB , $\bar{A}B$, $A\bar{B}$, $\bar{A}\bar{B}$

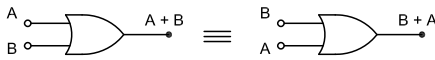
A product term is equal to 1 if and only if each of the literals in the term is one(1). A product term is equal to 0 when one or more of the literal are 0.

Laws of Boolean algebra

Commutative law

The commutative law for addition for two variables is written algebraically $A + B = B + A$ as shown in Fig 10a.

Fig 10a



The commutative law for two variable multiplication is

$AB = BA$ as shown in Fig 10b.

This law states that the order in which the variables are ORed/ANDed make no difference.

Fig 10b



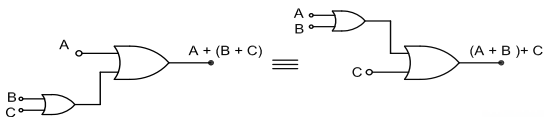
Associative law

The associative law of addition is written algebraically as follows for three variables as shown in Fig 10c.

$$A + (B + C) = (A + B) + C$$

The associative law of multiplication is written as follows for three variables.

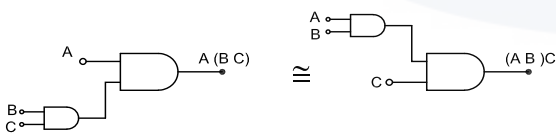
Fig 10c



$$A(BC) = (AB)C$$

This law states that it makes no difference in what order the variables are grouped when ORing/ANDing more than two variables.

Fig 10d



Distributive law

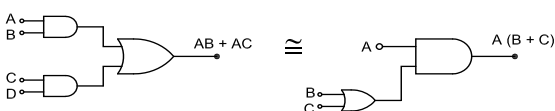
The distributive law is written for three variable as follows.

$$A(B + C) = AB + AC$$

This law states that ORing two or more variables and ANDing the result with a single variable is equivalent to ANDing the single variable with each of the two or more variables and then ORing the products as shown in Fig 10e. The distributive law also express the process of factoring in which the common variable 'A' is factored out of the product terms.

$$\text{Ex: } AB + AC = A(B + C)$$

Fig 10e



Boolean Algebra Rules

- 1 $A + 0 = A$
- 2 $A + 1 = 1$
- 3 $A + A = A$
- 4 $A + \bar{A} = 1$
- 5 $A + AB = A$
- 6 $A + \bar{A}B = A + B$
- 7 $A \cdot 0 = 0$
- 8 $A \cdot 1 = A$
- 9 $\bar{\bar{A}} = A$
- 10 $A \cdot A = A$
- 11 $A \cdot \bar{A} = 0$
- 12 $(A + B)(A + C) = A + BC$

De-Morgans theorem

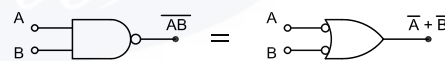
I theorem

The complement of a product of variables is equal to the sum of the complements of the variables.

$$\overline{AB} = \bar{A} + \bar{B}$$

The complement of two or more variables ANDed is equivalent to the OR of the complements of the individual variables. The related figure is shown in Fig 11a.

Fig 11a



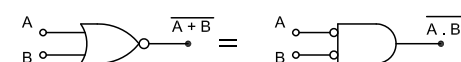
II theorem

The complement of a sum of variables is equal to the product of the complements of the variables.

$$\overline{A + B} = \bar{A} \cdot \bar{B}$$

The complement of two or more variables ORed is equivalent to the AND of the complements of the individual variables as shown in Fig 11b.

Fig 11b



Simplify the equation using De-Morgan's theorem

- 1 $(\overline{A+B+C})\bar{D} = \overline{A+B+C} \cdot \bar{D} \quad (\overline{AB} = \bar{A} + \bar{B})$
 $= \bar{A} \cdot \bar{B} \cdot \bar{C} + \bar{D} \quad (\overline{A+B} = \bar{A} \cdot \bar{B})$
- 2 $\overline{ABC+DEF} = \overline{ABC} \cdot \overline{DEF} \quad (\overline{A+B} = \bar{A} \cdot \bar{B})$
 $= (\bar{A} + \bar{B} + \bar{C}) \cdot (\bar{D} + \bar{E} + \bar{F})$

Simplification of Boolean equations

Prove that $A + \bar{A}B = A + B$

LHS

$$= A + \bar{A}B$$

$$= (A + AB) + \bar{A}B$$

$$= AA + AB + \bar{A} \cdot B$$

$$= AA + AB + \bar{A}B + 0$$

$$= AA + AB + \bar{A}B + A\bar{A}$$

$$= A(A + B) + \bar{A}(B + A)$$

$$= (A + \bar{A})(A + B)$$

$$= 1 \cdot (A + B)$$

$$= A + B$$

Since
 $A = A.A;$

$$A\bar{A} = 0;$$

Prove that $(A+B)(A+C) = A + BC$

$$\text{LHS} = (A+B)(A+C)$$

$$= AA + AB + AC + BC$$

$$= A + AC + AB + BC$$

$$= A(1+C) + AB + BC$$

$$= A + AB + BC$$

$$= A(1+B) + BC$$

$$= A + BC$$

Simplification of logic circuit using Boolean Equation

$$\bar{A}BC + A\bar{B}C + A\bar{B}\bar{C} + A\bar{B}C + A\bar{B}C + A\bar{B}C + A\bar{B}C + A\bar{B}C$$

I step: Simplify the Boolean equations

$$\bar{A}BC + \bar{B}C + AC$$

$$\bar{A}BC + \bar{B}C + AC$$

$$\bar{A}BC + AC + \bar{B}C$$

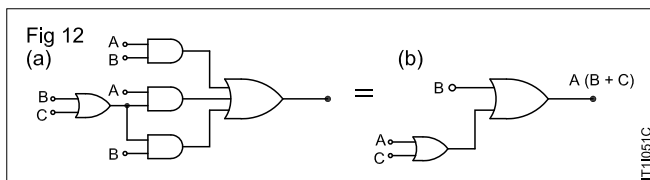
$$C(A + \bar{A}B) + \bar{B}C$$

$$C(A+B) + \bar{B}C$$

II step write logic diagram for the equations

$$AC + BC + \bar{B}C$$

Circuit before simplification is shown in Fig 12a. Circuit after simplification is shown in Fig 12b.



The above logic diagram and corresponding Boolean equations show how one can use Boolean Algebra for simplification of logic circuits for the desired logic output.

From the above example it is proved that how the logic circuit gates can be reduced for the same set of output result, using Boolean Algebra. The reduced logic circuit

consumes less power and propagation delay time is also reduced, in other words the speed of the circuit increases.

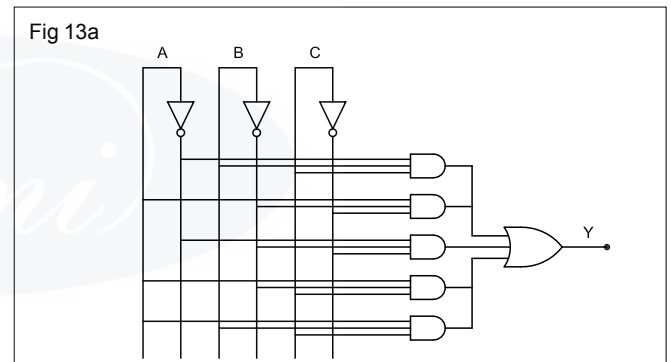
Example

- 2 Simplify the Boolean expression, and write logic diagram for the given equation and for simplified equation.

$$[\text{Since } A + \bar{A}B = A + B]$$

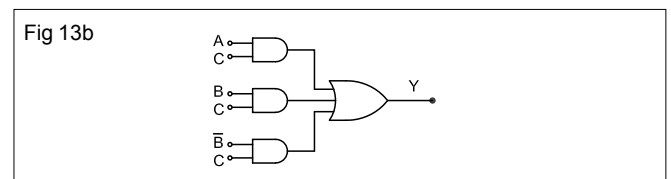
Logic diagram for the given equation is shown in Fig 13a.

$$\bar{A}BC + A\bar{B}C + \bar{A}\bar{B}C + A\bar{B}C + ABC = Y$$



Logic diagram for the simplified equation is shown in Fig 13b.

$$AC + BC + \bar{B}C = Y$$

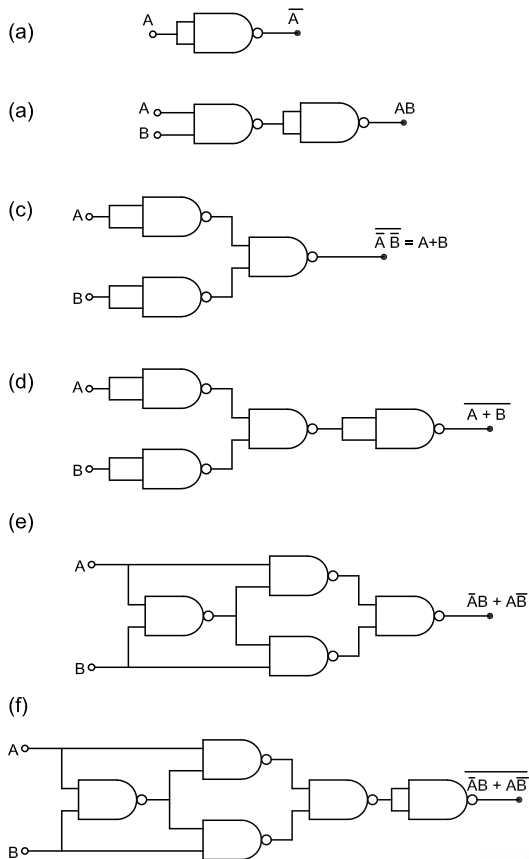


Universal property of NOR gate and NAND gate

The universality of the NAND or NOR gate means that it can be used as an inverter and that combination of NAND/NOR gate alone can be used to implement the all other remaining gates. (such as AND, OR, EXOR, EX-NOR)

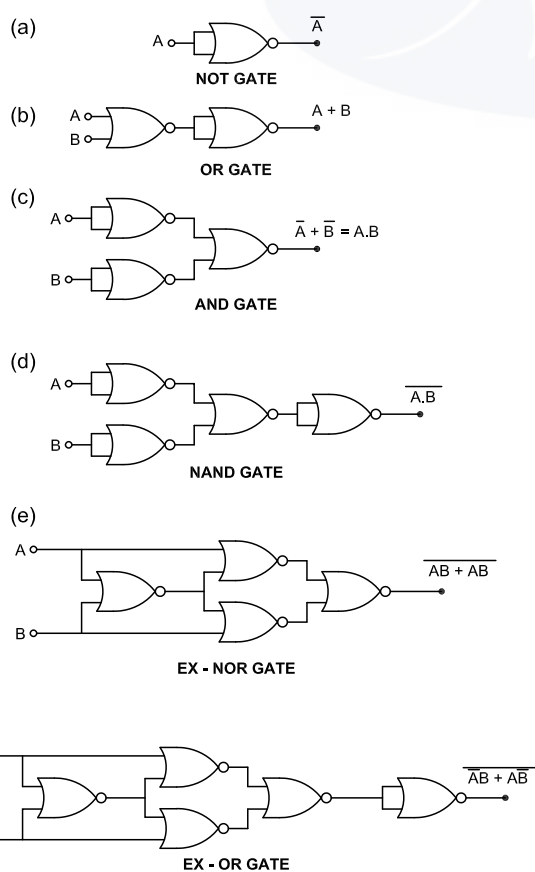
NAND gate as a universal logic element is shown in Fig 14a to Fig 14f.

Fig 14



Similarly NOR gate as universal logic gate as shown in Fig 15a to 15f.

Fig 15



Karnaugh Map

The Karnaugh map (K Map) is a graphical tool used to simplify a logic equation or to convert a truth table to its corresponding logic circuit in a simple orderly process. It is easier to simplify an expression using K-Maps when the number of variables is higher.

Truth table gives the value of output for each combination of input values, whereas the K Map gives the same information in a map.

Two types of methods are used for writing logic circuit expressions

Sum of products and Products of sums

Sum of products (SOP) expression consists of two or more AND terms that are ORed together.

Example of sum of products

$$ABC + \bar{A}'\bar{B}'C$$

Product of sum expression consists of two or more OR terms that are AND ed together.

Example of product of sum

$$(A + \bar{B} + C)(A + C)$$

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	1

The K-map drawn for the above truth table is given below in Fig.

	$\bar{B}'$	B 1
	0	0
A1	1	1

Steps for drawing the above K-map

Each case in the truth table corresponds to a square in the K-map.

- 1 A=0, B=0 condition in truth table corresponds to $\bar{A}'\bar{B}'$ square in the K-Map. The truth table shows Y=0 for A=0, B=0 so a '0' is placed in the $\bar{A}'\bar{B}'$ square in the K-Map.
- 2 A=0, B=1 condition in truth table corresponds to $\bar{A}'B$ square in the K-map. Because Y=1 for this case a '1' is place in the $\bar{A}'B$ square.
- 3 A=1, B=1 condition in truth table corresponds to AB square in the K-Map. Because Y=1 for this case a '1' is place in the AB square.

- 4 Similarly $A=1, B=0$ condition in truth table corresponds to $A\bar{B}$ square in the K-Map. Because $Y=1$ for this case a '1' is placed in the $A\bar{B}$ square.
- 5 The logic expression can be simplified by properly combining the squares in the K-Map that contain 1.
- 6 In Fig. This map contains a pair of 1's that are vertically and horizontally adjacent to each other. The first represents the $AB'+AB = A$ and the second represents $A'B+AB = B$. Combining together we get

$$Y = A + B$$

	B'	B
	0	1
A'0	0	0
A1	1	1

So far we have focussed on K-maps for three variables. Karnaugh Maps are useful for more than

three variables, and we'll look at how to extend ideas to four variables here. Shown below is a K-map for four variables.

Note the following about the four variable Karnaugh Map.

		YZ			
		00	01	11	10
WX	00	1	0	0	1
	01	1	1	0	0
	11	1	0	0	0
	10	1	0	0	1

- There are 16 cells in the map. Anytime you have N variables, you will have 2^N possible combinations, and 2^N places in a truth table or Karnaugh Map.
- Imagine moving around in the Karnaugh Map. Every time you cross a horizontal or vertical boundary one - and only one - variable changes value.
- The two pairs of variables - WX and YZ - both change in the same pattern.

Otherwise, if you can understand a Karnaugh Map for a three-variable function, you should be able to understand one for a four-variable function. Remember these basic rules that apply to Karnaugh maps of any size.

- In a Karnaugh Map of any size, crossing a vertical or horizontal cell boundary is a change of only one variable - no matter how many variables there are.
- Each single cell that contains a 1 represents a minterm in the function, and each minterm can be thought of as a "product" term with N variables.

- To combine variables, use groups of 2, 4, 8, etc. A group of 2 in an N-variable Karnaugh map will give you a "product" term with N-1 variables. A group of 4 will have N-2 variables, etc.
- You will never have a group of 3, a group of 5, etc. Don't even think about it. See the points above.

Let's look at some examples of groups in a 4-variable Karnaugh Map. Example 1 - A Group of 2

Here is a group of 2 in a 4-variable map.

		YZ			
		00	01	11	10
WX	00	0	0	0	0
	01	1	1	0	0
	11	0	0	0	0
	10	0	0	0	0

Note that Y and Z are 00 and 01 at the top of the two columns in which you find the two 1s. The variable, Z, changes from a 0 to a 1 as you move from the left cell to the right cell. Consequently, these two 1s are not dependent upon the value of Z, and Z will not appear in the product term that results when we combine the 1s in this group of 2. Conversely, W, X and Y will be in the product term. Notice that in the row in which the 1s appear, $W = 0$ and $X = 1$. Also, in the two columns in which the 1s appear we have $Y = 0$. That means that the term represented by these two cells is:

$$- W \cdot X \cdot Y$$

Larger groups in Karnaugh Maps of any size can lead to greater simplification. Let's consider the group shown shaded below. There are four terms covered by the shaded area.

		YZ			
		00	01	11	10
WX	00	1	0	0	1
	01	0	0	0	0
	11	0	0	0	0
	10	1	0	0	1

In the upper left:-

$$- \bar{W} \cdot \bar{X} \cdot \bar{Y} \cdot \bar{Z}$$

In the upper right;

$$- \bar{W} \cdot \bar{X} \cdot Y \cdot \bar{Z}$$

In the lower left;

$$- W \cdot \bar{X} \cdot \bar{Y} \cdot \bar{Z}$$

In the lower right;

$$- W \cdot \overline{X} \cdot Y \cdot \overline{Z}$$

These terms can be combined (assuming they are all ones in the Karnaugh Map!). The result is

By combining the first two terms above (the two terms at the top of the Karnaugh Map):-

$$- \overline{W} \cdot \overline{X} \cdot \overline{Z}$$

By combining the last two terms above (the two terms at the bottom of the Karnaugh Map):-

$$- W \cdot \overline{X} \cdot \overline{Z}$$

Then, these two terms can be combined to give:

$$- \overline{X} \cdot \overline{Z}$$

Notice how making the grouping larger reduces the number of variables in the resulting terms. That simplification helps when you start to connect gates to implement the function represented by a Karnaugh map.

By now you should have inferred the rules for getting the sum-of-products form from the Karnaugh map.

The number of ones in a group is a power of 2. That's 2, 4, 8 etc.

If a variable takes on both values (0 and 1) for different entries (1s) in the Karnaugh Map, that variable will not be in the sum-of-products form. Note that the variable should be one in half of the K-Map ones and it should be zero (inverted) in the other half.

If a variable is always 1 or always zero (it appears either inverted all the time in all entries that are one, or it is always not inverted) then that variable appears in that form in the sum-of-products form.

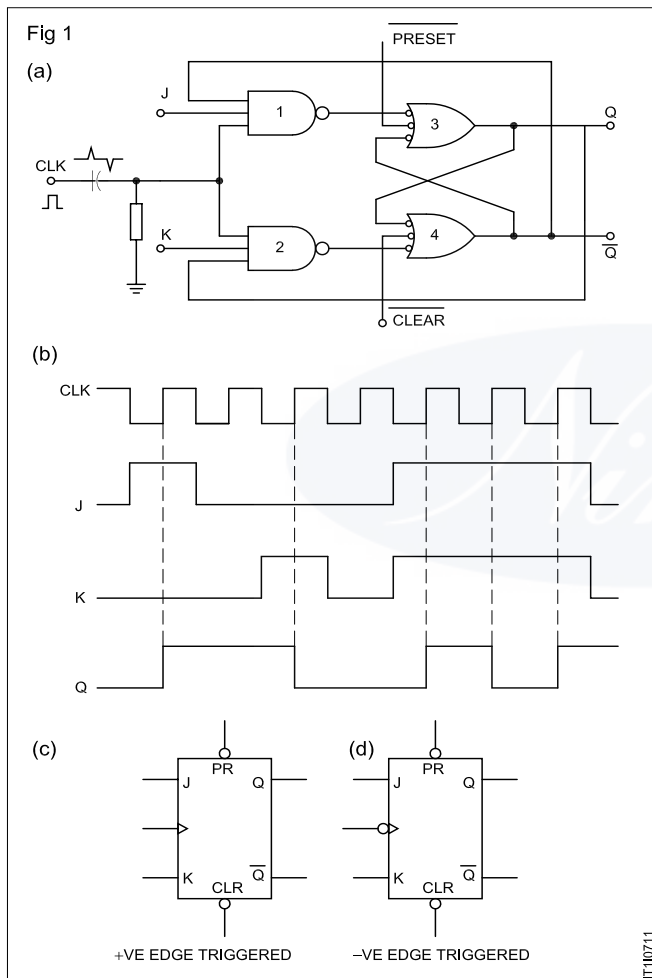


J.K Flip-flop

Objectives: At the end of this lesson you shall be able to

- explain construction of J.K flip-flop using NAND gates
- explain the function of Preset and clear inputs
- explain what is meant by active low and active high
- explain the working function of J.K master slave flip-flop
- explain flip-flop applications
- explain binary counter and their types
- explain asynchronous ÷ 16 counter using flip-flops
- explain modulus of a counter.

Operation of J-K flip flop



The figure shows one way to build a J.K flip-flop. The variables J and K are called control inputs. An R.C circuit with a short time constant, converts the rectangular clock pulse to narrow spikes. Because of the double inversion through the NAND gates, the circuit is +ve edge triggered. In other words, the input gates are enabled only on the rising edge of the clock as shown in truth table.

Reset

When J is low and K is high the upper output gate is disabled. So there is no way to set the flip flop. The only possibility is reset. When Q is high, the lower gate passes a reset trigger as soon as the +ve clock edge arrives. This forces Q to become low. Therefore J=0 and k=1 means that a rising clock edge resets the flip-flop.

Set

When J is high and K is low, the lower output gate is disabled. So it is impossible to reset the flip-flop. But flip-flop can be set, when Q is low, Q is high, the gate 1 passes a set trigger on the positive clock edge. This drives Q into the high state. That is J=1 and K=0 means that the next positive clock edge sets the flip flop.

Toggle

When J and K are both high, it is possible to set or reset the flip flop depending on the current state of the output if Q is high, the lower gate passes a reset trigger on the next positive clock edge on the other hand. When Q is low the upper gate passes a set trigger on the next positive clock edge. Either way Q changes to the complement of the last state. Therefore J=1 and K=1 means that the flip-flop will toggle on the next positive clock edge.

To summarize the operation of the J.K flip-flop, the circuit is inactive when the clock is low, high or on its -ve edge. Likewise the circuit is inactive when J and K are both low. Output changes occur only on the rising edge of the clock as indicated by the last three entries of the table. The o/p either resets, sets or toggles.

Truth table

CLK	J	K	Q	
0	x	x	NC	
—	x	x	NC	
—	x	x	NC	
x	0	0	NC	
-	0	1	0	Reset
-	1	0	1	Set
-	1	1	Toggle	Toggle

Racing

Toggling more than once during a clock cycle is called Racing. Assume that the circuit is level clocked. In other words, assume that RC circuit has been removed and run the clock straight, into the gates, with a high J, high K and high clock, the output toggles. New outputs are then fed back to the input gates. After two propagation times (input and output gates), the output toggles again. And once more new outputs return to the input gates. In this way the output can toggle repeatedly as long as the clock is high.

To overcome this racing problem, J.K master slave flip-flop has been developed.

Clear

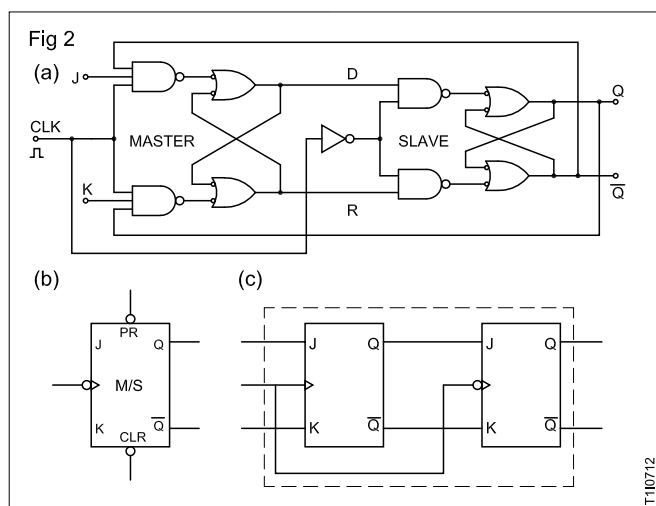
When power is first applied, flip-flops come up in random states. To get some computers started, an operator has to push a master reset button, this sends a clear (reset) signal to all flip-flops, normally clear signal will be active low, (i.e logic zero should be applied for clear the output. When clear is applied to gate-4 as shown in Fig 1a then the Q will be forced to Logic-0, then automatically Q will go to logic-1 condition. This signal, J and K signals have no control over output Q, when clear is set.

Pre-set

Like clear preset is an active low input. This input also independent of CLK, J & K inputs. When preset is made logic-0, the output Q is set to logic one. It is necessary in some digital system to preset the output before the system actually runs.

Master Slave Flip-flop

The figure shows the J.K.Master Slave Flip-flop. It provides another way to avoid racing. A master slave flip-flop is a combination of two clocked flip-flops connected in cascade. Master flip-flop is positive edge triggered, slave flip-flop is negative edge-triggered flip-flop.



- While the clock is high, the master is active and the slave is inactive.
- While the clock is low, the master is inactive and the slave is active.

The J.K master slave flip-flop is used as the main counting device. The popular IC 54LS/74LS76 is a dual JK master slave flip-flop.

Truth table for positive edge triggered JK flip flop

PR	CLR	CLK	J	K	Q
0	0	X	X	X	Race
0	1	X	X	X	1
1	0	X	X	X	0
1	1	X	0	0	NC
1	1	-	0	1	0
1	1	-	1	0	1
1	1	-	1	1	Toggle

Look at the Summarized truth table of J.K master slave flip-flop. A low PR and LOWCLR produces a race condition therefore, PR and CLR are normally kept at a high voltage when inactive. To clear, the flip flop make clear low, to preset the F/F make preset low.

Low J & Low K produces an inactive state regardless of the what the clock is doing. If K goes high by itself, the next clock pulse resets the flip-flop. If J goes high by itself, the next clock pulse sets the flip-flop when J & K are both high, each clock pulse toggle the state of flip flop.

Frequency division using flip flops

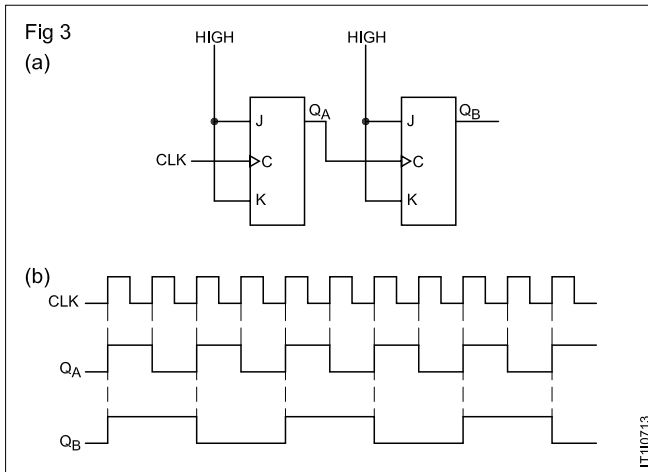
Flip-flops are used as frequency dividers of a periodic waveform. When a pulse waveform is applied to the clock input of a J.K flip-flop which is wired for toggle operation, provides square wave output with one half the frequency of the clock input. Thus a single flip-flop can be used for divide by -2 operation as illustrated in Fig 3. The flip-flop changes state on each triggering clock edge. This results in an output which is at half the frequency of the clock waveform.

Further division of clock frequency can be achieved by using the output of one flip-flop as the clock input to a second flip-flop as shown in Fig 3. The frequency of the QA output is divided by 2 by flip-flop B. The QB output is therefore, one fourth the frequency of the original clock input.

By connecting flip-flops in this way, a frequency division of 2^n is achieved, where n is the number of flip-flops. For example, three flip-flops divide the clock frequency by $2^3 = 8$. Four flip-flops divide the clock frequency by $2^4 = 16$; and so on.

Counter

A counter is one of the most useful and versatile subsystems in a digital system. A counter driven by a clock can be used to count the number of clock cycles. Since the clock pulses occur at known intervals, the counter can be used as an instrument for measuring time and therefore period or frequency.



Basically there are two types of counters

- 1 Synchronous counter
- 2 Asynchronous counter

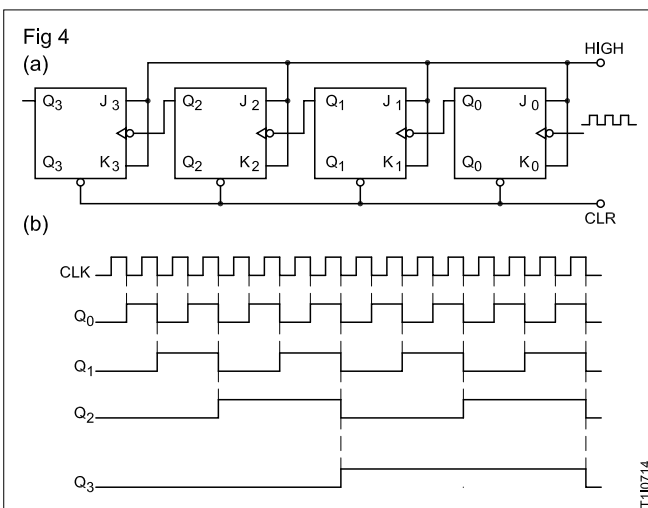
Synchronous counter

In this counter every flip-flop is triggered by the clock parallelly (synchronously) and thus settling time is simply equal to the delay time of a single flip-flop. The increase in speed is usually obtained at the price of increased hardware.

Asynchronous counter

Asynchronous counters are simple and straight forward in operation and construction and usually requires a minimum of hardware, however have a speed limitation. Each flip-flop is triggered by the previous flip-flop, [i.e., clock is applied serially] and thus the counter has a cumulative settling time. Counters such as these are also called as serial counters or ripple counters.

Serial and parallel counters are used in combination to compromise between speed of operation and hardware count. Serial, parallel or combination counters can be designed such that each clock pulse advances the contents of the counter by one, it is then operating in a count-up mode. The opposite is also possible; the counter then operates in the count-down mode. Furthermore, many counters can be either 'cleared' so that every flip-flop contains a zero, or preset such that the contents of the flip-flops represent any desired binary number.



Ripple counter

The Fig shows a counter built with J.K flip-flops since the J&K inputs are tied to a high voltage, each flip-flop will toggle, when its clock input receives a negative edge.

Visualize the Q outputs as a binary word.

$$Q = Q_3 Q_2 Q_1 Q_0$$

Q_3 is the most significant bit (MSB) and Q_0 is the least significant bit (LSB). When CLR goes low, all flip-flops reset, this results in a digital word of $Q = 0000$.

When clear returns to high, the counter is ready to go, since the LSB flip-flop receives each clock pulse, Q_0 toggles once per negative clock edge, as shown in Fig 4b. The remaining flip-flops toggle less often because they receive their negative edges from the preceding flip-flop outputs(Q). The triggers move through the flip-flops like a ripple in water, that is why this type of counter is called ripple counter.

Counting sequence

When CLR goes from low to high, the Q output of all flip-flops become zeroes.

$$Q = Q_3 Q_2 Q_1 Q_0$$

$$Q = 0000$$

When the first clock pulse triggered the LSB flip-flop, its output Q_0 becomes 1, so the first output word is $Q=0001$.

When the second clock pulse arrives, Q_0 resets to zero, and this negative falling edge sets next flip-flop output Q_1 to 1 as shown in the timing diagram. Therefore, the next output word is $Q=0010$.

The third clock pulse advances Q_0 to 1, at this stage Q_0 is having +ve rising edge, so Q_1 output will not change then the counter output is $Q=0011$.

The fourth clock pulse forces the Q_0 flip-flop to reset and carry. In turn, the Q_1 flip-flop resets and carry, in turn, the Q_2 flip-flop resets and carry.

The resulting output word is $Q = 0100$.

The fifth clock pulse gives $Q = 0101$.

The sixth clock pulse gives $Q = 0110$.

and the seventh pulse gives $Q = 0111$.

ON the eighth clock pulse Q_0 resets and carrier, Q_1 resets and carrier, Q_2 resets and carry, and Q_3 advances to 1. So the output word becomes $Q = 1000$.

The ninth clock pulses gives $Q = 1001$.

The tenth gives $Q = 1010$ and so on

at the 15th clock pulse $Q = 1111$

The 16th clock pulse resets all flip-flops. Therefore the counter resets to $Q = 0000$ and the cycle repeats.

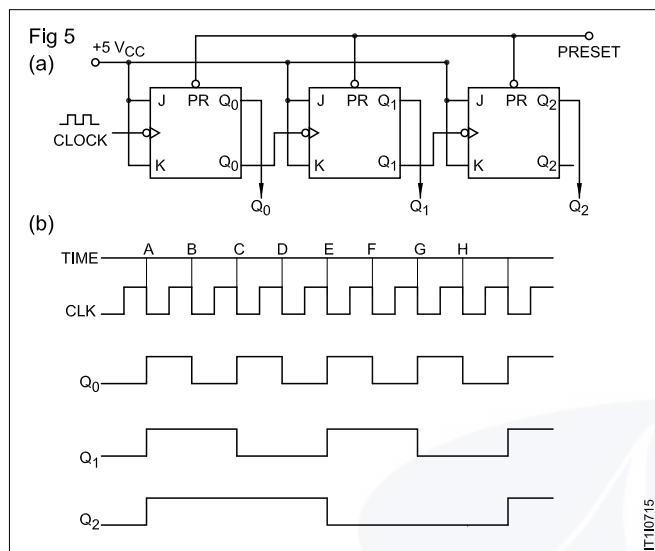
By adding more flip-flops to the left end of the counter circuit, the counter length can be extended. Eight flip-flops gives an 8 bit ripple counter, twelve flip-flops result in a 12 bit ripple counter and so on.

Down counter

The counter discussed above is an up counter, counts from 0 to 15 [0000 to 1111]. If a counter counts from 1111 to 0000 then it is called DOWN counter. In a down counter Q_0 output as shown in Fig 5 is connected to the clock input of the next flip-flop. Each flip-flop toggles when its clock input goes from 1 to 0. Flip flop Q_0 toggles with each negative clock transition as before. But flip flop Q_1 will toggle each time Q_0 goes high. Notice that each time Q_0 goes high, Q_0 goes low, and it is this negative transition on Q_0 that triggers Q_1 .

A low preset signal sets all output, producing an output word of $Q = 1111$.

When pre goes high, the action starts.



The first clock pulse produces a negative toggle in Q_0 , nothing else happens. $Q = 1110$

The second clock pulse produces a positive toggle in Q_0 , which produces a negative toggle in Q_1 . $Q = 1101$

On the third clock pulse Q_0 toggles negatively and $Q = 1100$

On the fourth clock pulse, Q_0 toggles positively Q_1 toggles positively and Q_2 toggles negatively. $Q = 1011$.

Likewise counting down from 15-0 takes place, when count reaches 0, i.e. $Q=0000$, on the next clock pulse, all flip-flops toggles positively to get $Q=1111$ and the cycle repeats.

Modulo -10 Counter

Asynchronous Decade counters

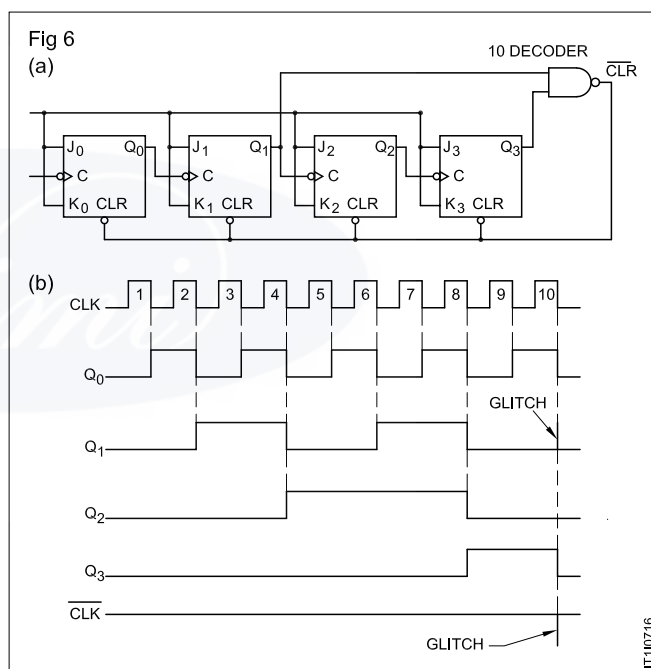
Regular binary counters have a maximum modulus, which means they progress through all of their possible states. The maximum possible number of states (maximum modulus) of a counter is 2^n , where n is the number of flip-flops in the counter.

Counters can also be designed to have a number of states in their sequence that is less than the maximum of 2^n . The resulting sequence is called a truncated sequence.

One common modulus for counters with truncated sequence is ten. Counters with ten states in their sequence (modulus-10) are called decade counters. A decade counter with a count sequence of zero (0000) through nine (1001) is a BCD decade counter because its ten-state sequence is the BCD code. This type of counter is useful in display applications in which BCD is required for conversion to a decimal readout.

To obtain a truncated sequence, it is necessary to force the counter to recycle before going through all of its normal states. For example, the BCD decade counter must recycle back to the 0000 state after the 1001 state. A decade counter requires four flip-flops (three flip-flops are insufficient because $2^3 = 8$).

We will use a 4-bit asynchronous counter such as the one in Fig 6a and modify its sequence to illustrate the principle of truncated counters. One way to make the counter recycle after the count of nine (1001) is to decode count ten (1010) with a NAND gate and connect the output of the NAND gate to the clear () inputs of the flip-flops, as shown in Fig 6.



Notice in Fig 6a that only Q_1 and Q_3 are connected to the NAND gate inputs. This arrangement is an example of partial decoding, in which the two unique states ($Q_1=1$ and $Q_3=1$) are sufficient to decode the count of ten, because none of the other states (zero through nine) have both Q_1 and Q_3 HIGH at the same time. When the counter goes into count ten (1010), the decoding gate output goes LOW and asynchronously resets all the flip-flops.

The resulting timing diagram is shown in Fig 6b. Notice that there is a glitch on the Q_1 waveform. The reason for this glitch is that Q_1 must first go HIGH before the count of ten can be decoded. Not until several nanoseconds after the counter goes to the count of ten does the output of the decoding gate go LOW (both inputs are HIGH). Thus, the counter is in the 1010 state for a short time before it is reset to 0000, thus producing the glitch on Q_1 and the resulting glitch on the CLR line which resets the counter.

Other truncated sequences can be implemented in a similar way.

Modulo-12 counter

An asynchronous counter can be implemented having a modulus of twelve with a straight binary sequence from 0000 through 1011.

Since three flip-flops can produce a maximum of eight states, four flip-flops are required to produce any modulus greater than eight but less than or equal to sixteen.

When the counter gets to its last state, 1011, it must recycle back to 0000 rather than going to its normal next state of 1100, as illustrated in the following sequence chart:

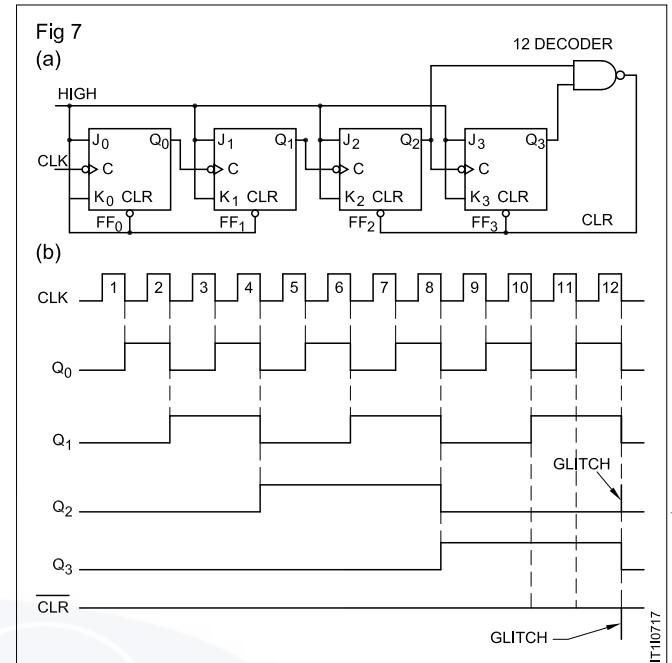
Q_3	Q_2	Q_1	Q_0
0	0	0	0
.	.	.	.
.	.	.	.
1	0	1	1
1	1	0	0

← Recycles

Normal next state

←

Observe that Q_0 and Q_1 both go to 0 anyway, but Q_2 and Q_3 must be forced to 0 on the twelve clock pulse. Fig 7a shows the modulus-12 counter. The NAND gate partially decodes count twelve (1100) and resets flip-flop 2 and flip-flop 3. Thus, on the twelve clock pulse, the counter is forced to recycle from count eleven to count zero, as shown in the timing diagram of Fig 7b. (It is in count twelve for only a few nanoseconds before it is reset by the glitch on CLR.)



Up/down counters

Objectives: At the end of this lesson you shall be able to

- explain the basic operation of an up/down counter
- discuss the 74LS190 up/down decade counter.

UP/DOWN SYNCHRONOUS COUNTERS

An up/down counter is one that is capable of progressing in either direction through a certain sequence. An up/down counter, sometimes called a bidirectional counter, can have any specified sequence of states. A 3-bit binary counter that advances upward through its sequence (0,1,2,3,4,5,6,7) and then can be reversed so that it goes through the sequence in the opposite direction (7,6,5,4,3,2,1,0) is an illustration of up/down sequential operation.

In general, most up/down counters can be reversed at any point in their sequence. For instance, the 3-bit binary counter can be made to go through the following sequence:

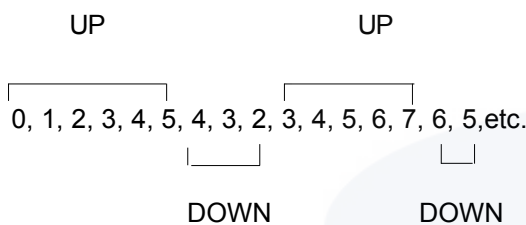
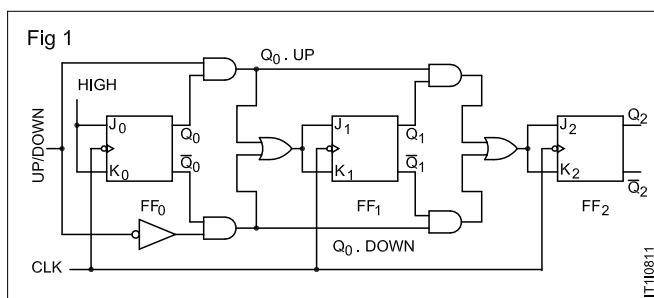


Table 1 shows the complete up/down sequence for a 3-bit binary counter. The arrows indicate the state-to-state movement of the counter for both its UP and its DOWN modes of operation. An examination of Q_0 of Fig 1 for both the up and down sequences shows that FFO toggles on each clock pulse. So the J_0 and K_0 inputs of FFO are

$$J_0 = K_0 = 1$$



For the up sequence, Q_1 changes state on the next clock pulse when $Q_0 = 1$. For the down sequence, Q_1 changes on the next clock pulse when $Q_0 = 0$. Thus, the J_1 and K_1 inputs of FFI must equal 1 under the conditions expressed by the following equation:

$$J_1 = K_1 = (Q_0 \cdot UP) + (Q_0 \cdot DOWN)$$

For the up sequence, Q_2 changes state on the next clock pulse when $Q_0 = Q_1 = 1$. For the down sequence, Q_2 changes on the next clock pulse when $Q_0 = Q_1 = 0$. Thus, the J_2 and K_2 inputs of FF2 must equal Q_1 under the conditions expressed by the following equation:

$$J_2 = K_2 = (Q_0 \cdot Q_1 \cdot \text{UP}) + (Q_0 \cdot Q_1 \cdot \text{DOWN})$$

Each of the conditions for the J and K inputs of each flip-flop produces a toggle at the appropriate point in the counter sequence.

A synchronous 4-bit binary up/down counter

A timing diagram and the sequence of a synchronous 4-bit binary up/down counter if the clock and UP/DOWN control inputs have waveforms as shown in Fig 2. The counter starts in the all 0s state and is positive edge-triggered.

Table 1

Clock pulse	UP	Q_2	Q_1	Q_0	Down
0		0	0	0	
1		0	0	1	
2		0	1	0	
3		0	1	1	
4		1	0	0	
5		1	0	1	
6		1	1	0	
7		1	1	1	

The timing diagram showing the Q outputs is shown in Fig 2b. From these waveforms, the counter sequence is as shown in Table 2.

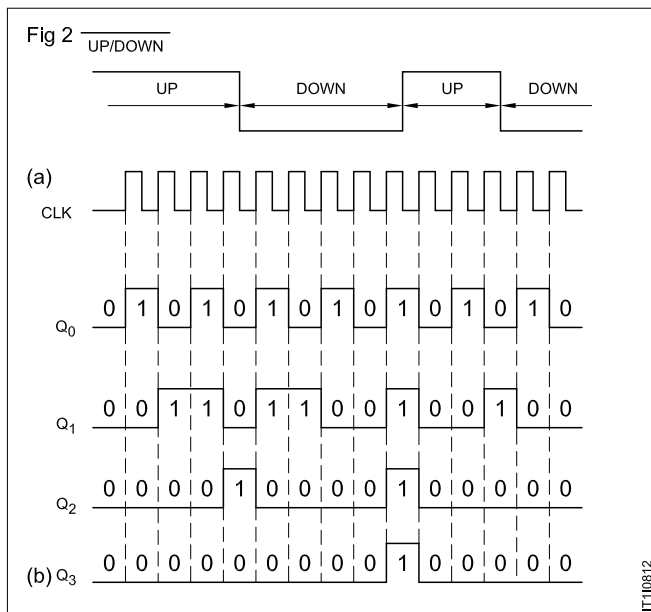
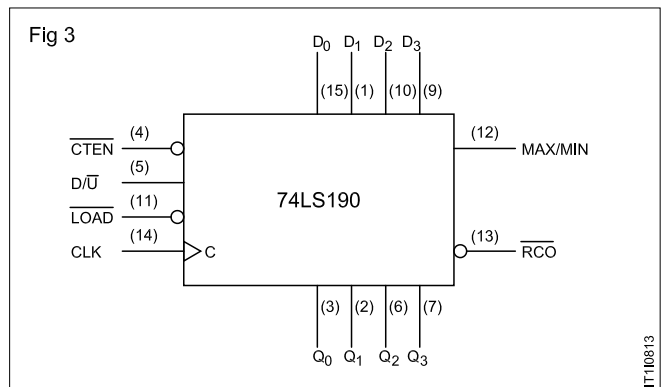


Table 2

Q ₃	Q ₂	Q ₁	Q ₀	
0	0	0	0	UP
0	0	0	1	
0	0	1	0	
0	0	1	1	
0	1	0	0	DOWN
0	0	1	1	
0	0	1	0	
0	0	0	1	
1	1	1	1	UP
0	0	0	0	
0	0	0	1	
0	0	1	0	
0	0	0	1	DOWN
0	0	0	0	

The 74LS190 Up/down decade counter

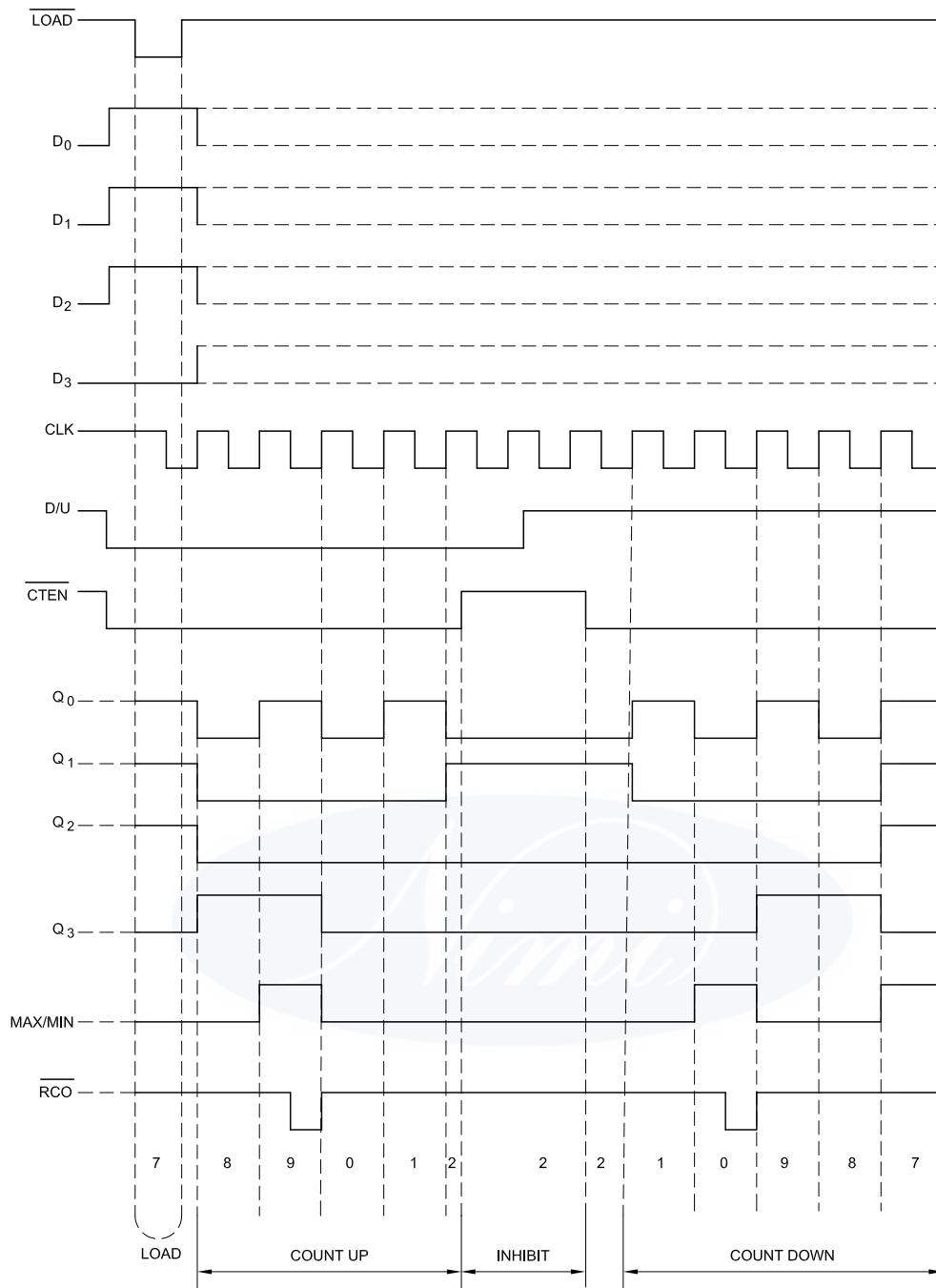
Fig 3 shows a logic diagram for the 74LS190, a good example of an integrated circuit up/down counter. The direction of the count is determined by the level of the up/down input ($\overline{D/U}$). When this input is HIGH, the counter counts down; when it is LOW, the counter counts up. Also, this device can be preset to any desired BCD digit as determined by the states of the data inputs when the LOAD input is LOW.



The MAX/MIN output produces a HIGH pulse when the terminal count nine (1001) is reached in the UP mode or when the terminal count zero (0000) is reached in the DOWN mode. This MAX/MIN output, along with the ripple clock output (RCO) and the count enable input (CTEN) is used when cascading counters.

Fig 4 is an example timing diagram showing the 74LS190 counter preset to seven (0111) and then going through a count-up sequence followed by a count-down sequence. The MAX/MIN output is HIGH when the counter is in either the all-0s state (MIN) or the 1001 state (MAX).

Fig 4



IT10814

Binary arithmetic

Objectives: At the end of this lesson you shall be able to

- explain binary arithmetic
- perform binary addition
- perform binary subtraction using 1's complement and 2's complement
- explain half adder circuit, full adder circuit
- explain the IC 74LS83, 4 bit parallel adder
- explain IC 74LS83 4 bit parallel adder can be used for subtraction.

Binary arithmetic is essential in all digital computers and in many other types of digital systems. To understand digital systems, you must know the basics of binary addition, subtraction, multiplication and division.

Binary addition

Physical quantities are represented by numbers. Addition represents combining of physical quantities. Digital computers do not process decimal numbers, they process binary numbers. Addition is a key process to perform subtraction, multiplication and division. The four basic cases for adding binary digits are as follows.

$0 + 0 = 0$; Sum is 0 with a carry of 0.

$0 + 1 = 1$; Sum is 1 with a carry of 0.

$1 + 0 = 1$; Sum is 1 with a carry of 0.

$1 + 1 = 10$; Sum is 0 with a carry of 1.

Notice that the first three cases result in a single bit and in the fourth case the addition of two 1's yields a binary two i.e. 10. When binary numbers are added, the last condition creates a sum of 0 in a given column and a carry of 1 over to the next column to the left, as illustrated in the following addition of $11 + 01$.

carry	1	1	
	0	1	1
+	0	0	1
	1	0	0

In the right most column, $1+1=0$ with a carry of 1 to the next left column. In the middle column, $1+1+0=0$ with a carry of 1(one) to the next left column. In the left most column, 1 remains as final carry of the 2 bit addition. Hence the result is 100.

Example:

1	carry	1110	
		14	1110
		10	1010
		24	11000

$$2 \quad 10 + 12$$

10		1010
12	+	1100
22		10110

The above process is column-by-column addition which can be applied to find the sum of two binary numbers of any length. The following example shows 8-bit arithmetic addition operation.

A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0
B_7	B_6	B_5	B_4	B_3	B_2	B_1	B_0
?							

The most significant bit (MSB) of each number is on the left side and least significant bit is on the right side. For the first number, A_7 is the MSB and A_0 is the LSB, similarly for the 2nd number B_7 and B_0 are the MSB and LSB respectively.

Signed numbers

Digital systems such as the computer, must be able to handle both +ve and -ve numbers, A signed binary numbers consists of both sign and magnitude information. The sign indicates whether a number is +ve or -ve and the magnitude is the value number. There are three ways in which signed numbers can be represented in binary form: sign magnitudes, 1s complement, and 2's complement.

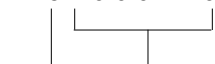
Sign-magnitude system

The left most bit in a signed binary number is the sign bit, which tells you whether the number is +ve or -ve, A zero in the left most position represents +ve number and a ONE represents -ve number. The remaining bits are the magnitude bits. The magnitude bits are in true (uncomplimented) binary form for both +ve and -ve numbers.

Example:

+25 is expressed as an 8 bit signed binary number using the sign magnitude system as

$$+25 = 00011001$$



Sign bit Magnitude bit

$$-25 = 10011001$$

Notice that the only difference between +25 and -25 is with the sign bit because the magnitude bits are same for both +ve and -ve numbers.

"In the sign-magnitude system, a -ve number has the same magnitude bits as the corresponding +ve number but the sign bit is a 1." Although sign magnitude system is straight forward, calculators and computers do not use it, because circuit implementation is more complex than other systems.

1's complement system

Positive numbers in the 1's complement system are represented the same way as the positive sign magnitude numbers. In the 1's complement system, a negative number is the 1's complement of the corresponding +ve number.

Example:

The decimal number -25 is expressed as the 1's complement of +25 (00011001) as 11100110.

i.e 1's complement of 00011001 (+25) = 11100110 (-25)

(The 1's complement of a binary number is obtained by simply changing each 0 to a 1 and each 1 to a 0).

Example:

Determine the decimal value of the signed binary numbers expressed in 1's complement.

$$11101000$$

The bits and their powers of two weights for the -ve number are as follows.

Notice that the -ve sign bit has a weight of -2^7 or -128.

-2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
1	1	1	0	1	0	0	0

Summing the weights where there are 1s.

$$1 \times -2^7 (128) = -128$$

$$1 \times 2^6 (64) = +64$$

$$1 \times 2^5 (32) = +32$$

$$1 \times 2^3 (8) = +8$$

$$= -128 + 104$$

$$= -24$$

Adding 1 to the result, the final number is $= -24 + 1 = -23$.

The decimal value of the signed number 11101000 expressed in 1's complement is $= -23$.

2's complement system

Positive numbers in the 2's complement system are also represented the same way as in sign magnitude and 1's complement system. Negative numbers are 2's complement of the corresponding positive no's.

2's complement of a binary number is found by adding ONE(1) to the LSB of the 1's complement.

$$2's \text{ Complement} = (1's \text{ complement}) + 1$$

Example:

Find the 2's complement of 1011011

Solution

$$1011011 \quad - \quad \text{Binary number}$$

$$0100100 \quad - \quad 1's \text{ complement}$$

$$1 \quad - \quad \text{Add 1}$$

$$+ 0100101$$

For example, the decimal number -25 can be expressed in binary form by writing 2's complement for +25.

$$+25 = 00011001 \quad - \quad \text{Binary number}$$

$$11100111 \quad - \quad 2's \text{ complement}$$

Example:

Express the decimal -39 as an 8 bit number in sign-magnitude using 2's complement system.

Solution

In the 2's complement system, -39 is produced by taking the 2's complement of +39 (00100111) as follows.

$$+39 = 00100111 \quad \text{Binary number}$$

$$11011000 \quad 1's \text{ complement}$$

$$+ \quad 1$$

$$11011001 \quad 2's \text{ complement}$$

$$-39 = 11011001$$

Verification =

-2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0
-128	64	32	16	8	4	2	1

$$1 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 1$$

Since the MSB of the binary equivalent is one so, 2^7 should be taken as -ve sign.

$$-128 + 64 + 16 + 8 + 1 = -128 + 89$$

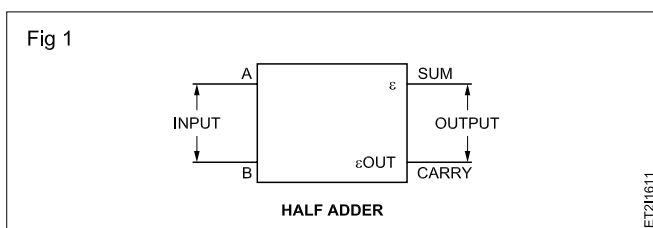
$$= -39$$

The 2's complement system is preferred for representing signed numbers as it requires a summation weights regardless of whether the number is +ve or -ve. It is used in most computers because it makes arithmetic operations easier.

Basic Adder

Adders are used in many types of digital systems in which numerical data are processed. Computers and calculators perform binary operations on two binary numbers at a time, where each number can have several binary digits. The logic symbol for a half adder is shown in Fig 1. There are two basic categories of adders.

- 1 Halfadder
- 2 Full adder



Half adder

The half-adder accept two binary digits on its inputs and produces two binary digits on its outputs, a sum bit and a carry bit.

Table 1 (Truth table)

A	B	Sum $S = A \oplus B$	Carry $C_{out} = AB$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

From the logic operation of the half-adder as stated in the Table 1, expression can be derived for the sum and the output carry as functions of the inputs, notice that the output carry is a 1 only when both A and B are 1s. Therefore carry (C_{out}) can be expressed as the AND of the input variables.

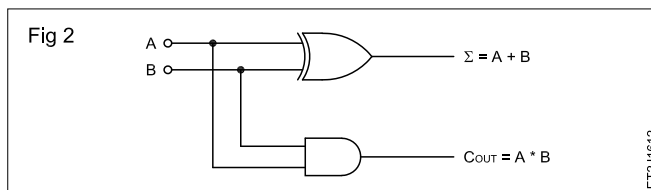
$$C_{out} = A.B \quad \longrightarrow 1$$

The sum output(S) is a 1 only if the input variables, A and B are not equal.

The sum can therefore be expressed as the exclusive -OR of the input variables.

$$\text{Sum}(S) = A \oplus B \quad \longrightarrow 2$$

From equation 1 and 2 the logic implementation required for the half-adder function can be developed. The output carry is produced with an AND gate with 'A' and 'B' on the inputs and the sum outputs is generated with an Ex-OR gate, as shown in Fig 2.



Full adder

The full adder accepts three inputs including an input carry and generates a sum output and an output carry.

The basic difference between a full-adder and a half-adder is that the full-adder accepts an input carry. A logic symbol for a full-adder is shown in Fig 3 and the truth table in the Table 2 shows the operation of a full-adder.

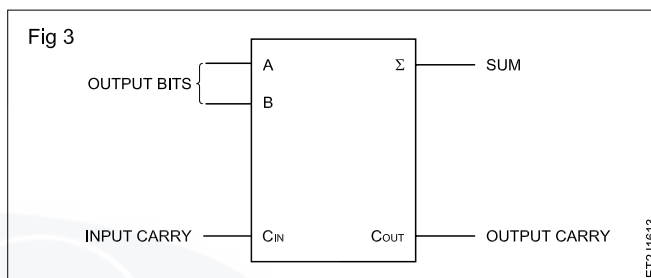


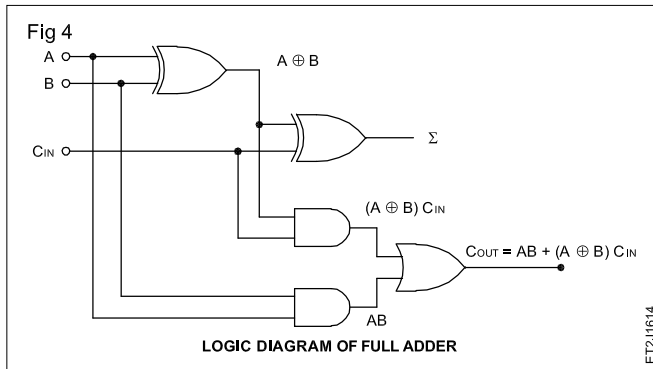
Table 2

A	B	C_{in}	Cout	S
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

The full-adder must add the two input bits along with the input carry. From the truth-table of the half-adder we know that the sum of the input bits A and B is $A \oplus B$. To get the sum output of the full adder the input carry (C_{in}) must be exclusive-ORed with $A \oplus B$. Then the sum

$$S = (A \oplus B) \oplus C_{in}$$

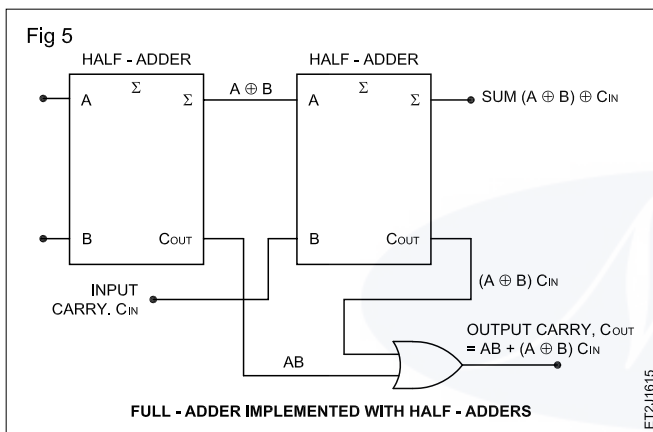
This means that to implement the full-adder sum function, two exclusive-OR gates can be used. The first must generate the term $A \oplus B$, and the second has the inputs from the output of the XOR gate and the input carry, as shown in Fig 4.



The output carry of the full-adder is therefore produced by the inputs A, ANDed with B and $A \oplus B$ ANDed with C_{in} . These two terms are ORed, and expressed in equation shown below and this function is implemented and combined with the sum logic to form a complete full-adder circuits, as shown in Fig 4.

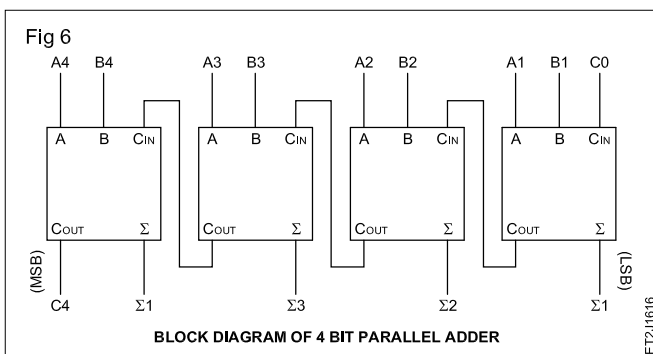
$$C_{out} = AB + (A \oplus B) C_{in}$$

The Fig 5, shows there are two half-adders, connected as shown in block diagram to form full-adder.



Four bit parallel adder

A basic 4-bit parallel adder is implemented with four full-adders as shown in the Fig 6.



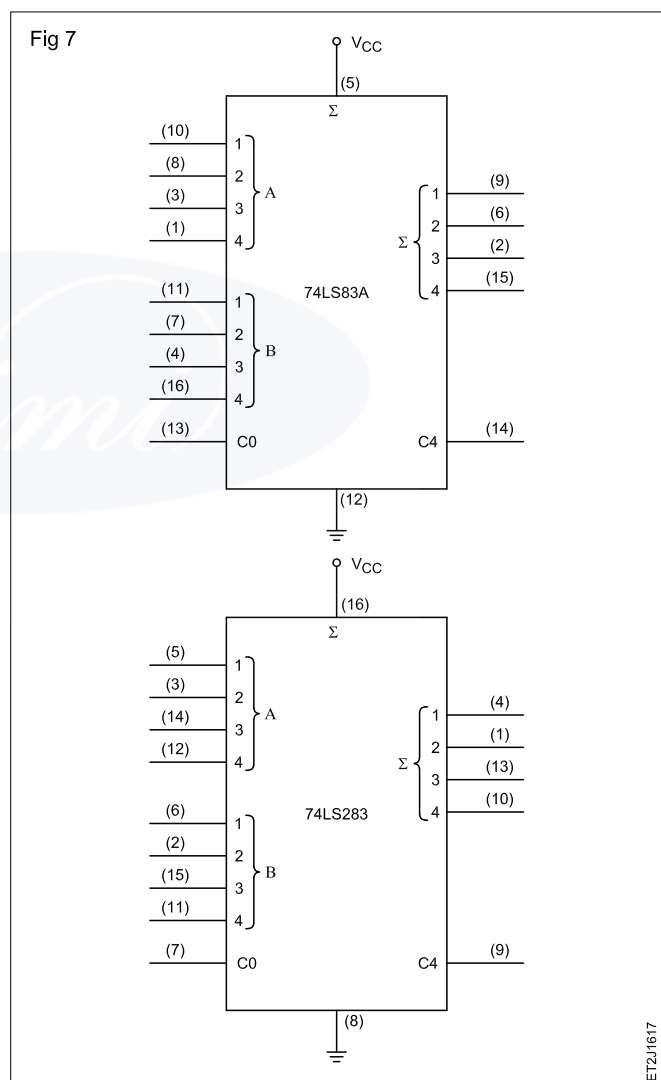
Block diagram of 4 bit parallel adder

The LSB, (A₁ and B₁) in each number being added into the right most full-adder; the higher order bits are applied as shown to the successively higher order adders, with MSBs (A₄ and B₄) in each number being applied to the left most full adder. The carry output of each adder is connected to the carry input of the next higher order adder as indicated.

In the manufacturer's data sheets the input labeled C_0 is the input carry to the least significant bit adder, C_4 is the output carry of the most significant bit adder, and S_1 (LSB) through S_4 (MSB) are the sum outputs.

74LS83 4 bit parallel adder

4-bit parallel adders that are available as Medium-Scale Integrated (MSI) circuits are the 74LS83A and the 74LS283 low-power Schottky TTL devices. These devices are also available in other logic families such as standard TTL (7483A and 74283) and CMOS (74HC283). The 74LS83A and the 74LS283 are functionally identical to each other but not pin compatible, that is the pin numbers for the inputs and outputs are different due to different power and ground pin connections. For the 74LS83A, V_{CC} is pin 5 and ground is pin 12 on the 16-pin package. For the 74LS283, V_{CC} is pin 16 and ground is pin-8, which is a more standard configuration. Logic symbols for both of these devices are shown in Fig 7 with pin numbers in parenthesis.



The 4 bit parallel adder can be expanded to handle the addition of higher bit numbers by a process called cascading. In this process, the carry output of the lower-order adder is connected to carry input of the higher-order adder being cascaded.

Binary subtraction

Subtraction is a special case of addition. For example Subtracting +6 (the subtrahend) from +9 (the minuend) is equivalent to adding -6 to +9. Basically the subtraction operation changes the sign of the subtrahend and adds it to the minuend. The result of a subtraction is called the difference.

$$9 - 6 = 9 + (-6)$$

The sign of a positive or negative binary number is changed by taking its 2's complement.

Example:

The result of 2's complement of the positive number 0110(+6) is 1's complement of the number + 1

$$\text{i.e. } 1001 + 1 = 1010$$

1010 is 2's complement of 0110(+6), which is equal to - 6 in decimal system, as shown below.

$$1 \quad 0 \quad 1 \quad 0$$

$$-8 + 0 + 2 + 0 = -6$$

Example:

Subtract 6 from 9 in 2's complement method

$$9 - 6 = 3 \quad \text{normal method}$$

$$9 + (-6) = 3 \quad \text{2's complement method}$$

Binary form

$$9 = 1001 \text{ (minuend)} \quad 1001 - 0110 = 0011$$

$$6 = 0110 \text{ (subtrahend)}$$

2's complement method

I step: 2's complement of subtrahend 0110 is

1's complement of subtrahend + 1

$$\text{i.e. } 1001 + 1 = 1010 \text{ (equal to -6 in decimal system)}$$

II step: Add minuend with 2's complement of subtrahend

$$\text{i.e. } 1001 + 1010 = 10011$$

Discard the carry 1, then the result is 0011.

Adder-Subtractor

Full adder can be used for to add or subtract binary numbers. The Fig 8 shows the how adder circuit can be used as subtractor.

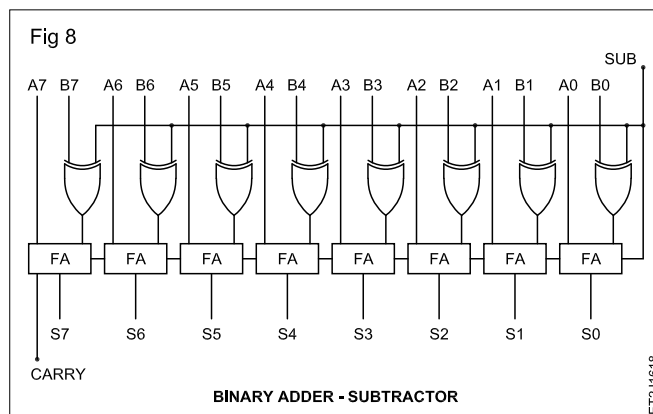
The carry out from each full adder is the carry into the next-higher full adder. The numbers being proposed are $A_7 \dots A_0$ and $B_7 \dots B_0$ while the output sum is $S_7 \dots S_0$. With 8 bit arithmetic, the final carry is ignored. With 16-bit arithmetic, the final carry is the carry into the addition of the upper byte.

Addition

$$A_7 \dots A_0$$

$$+ B_7 \dots B_0$$

$$S_7 \dots S_0$$



During an addition, the SUB signal is deliberately kept in the low state, therefore the binary number $B_7 \dots B_0$ passes through the controlled inverter (through Ex-OR gate) with no change. The full-adders then produce the correct output SUM.

For instance, suppose that the numbers being added are +125 and -67, then $A_7 \dots A_0 = 01111101$ and $B_7 \dots B_0 = 10111101$.

Since SUB=0 during an addition, the CARRY IN to the LSB column is 0.

$$\begin{array}{r} 0 \quad \leftarrow \text{Sub input} \\ 01111101 \quad \leftarrow \text{Input 'A'} \\ + 10111101 \quad \leftarrow \text{Input 'B'} \\ \hline 100111010 \end{array}$$

During 8 bit arithmetic operation 'last carry' is ignored, therefore the answer is $S_7 \dots S_0 = 00111010$.

Subtraction

$$\begin{array}{r} A_7 \dots A_0 \quad \text{(minuend)} \\ (-) B_7 \dots B_0 \quad \text{(subtrahend)} \\ \hline S_7 \dots S_0 \end{array}$$

During the subtraction, the SUB signal is deliberately put into high state. Therefore the controlled inverter (Ex-OR gates) produces the 1's complement of 'B' inputs, because the SUB is the carry IN, to the first full-adder (tied to logic-1) circuit processes the data as given.

$$\begin{array}{r} 1 \quad \leftarrow \text{SUB} \\ A_7 \quad A_6 \quad A_5 \quad A_4 \quad A_3 \quad A_2 \quad A_1 \quad A_0 \\ + \quad B_7 \quad B_6 \quad B_5 \quad B_4 \quad B_3 \quad B_2 \quad B_1 \quad B_0 \\ \hline S_7 \quad S_6 \quad S_5 \quad S_4 \quad S_3 \quad S_2 \quad S_1 \quad S_0 \end{array}$$

When $A_7 \dots A_0$ is applied with all zeros the circuit produces the 2's complement of $B_7 \dots B_0$ because 1 is being added to the 1's complement $B_7 \dots B_0$, when $A_7 \dots A_0$ doesn't equal zero the effect is equivalent to adding $A_7 \dots A_0$ and the 2's complement of $B_7 \dots B_0$.

Example:

(A - B)

82 - 17

A B S

= 01010010 - 00010001 = ?

The controlled inverter produces the 1's complement of B, which is 11101110, since SUB=1 during a subtraction, the circuit performs the following condition.

1	<— SUB
01010010	<— A input
11101110	<— B input
101000001	<— S output

For 8-bit arithmetic, the final carry is ignored, therefore the answer is $S_7 \dots S_0 = 01000001$.

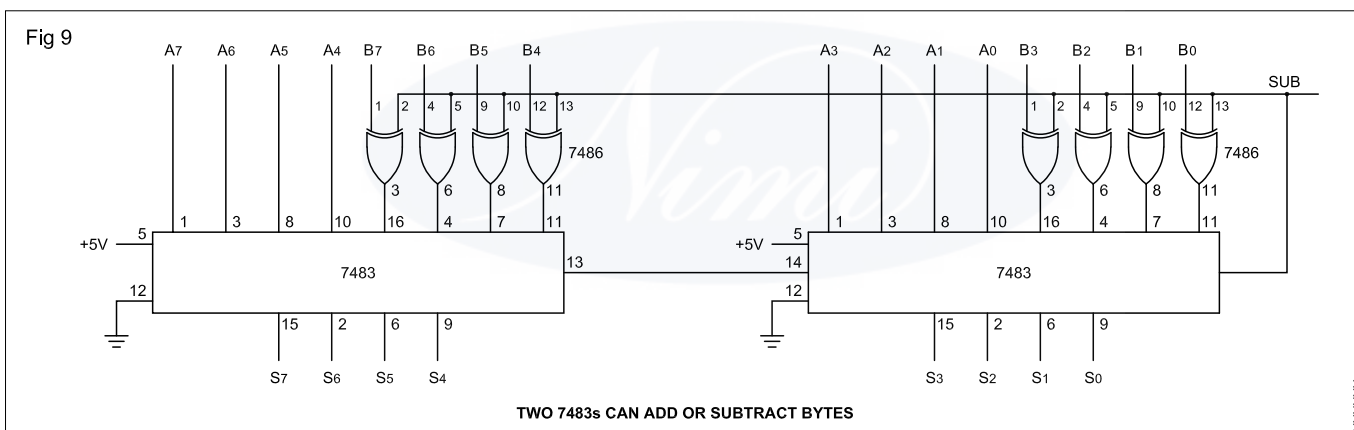
This answer is equivalent to decimal +65 which is the algebraic difference between the number +82 and +17.

Subtraction circuit based on 2's complement method using Adder IC

In this circuit the Minuend is applied to the A inputs of IC 7483 and Subtrahend is fed to the B inputs, through EX-OR gates and output is taken at S outputs, as shown in Fig 9.

The IC 7486 is an Exclusive-OR gate used for 1's complementing the SUBTRAHEND.

Carry input and one input from each Ex-or-gate is tied to some logic status. For addition SUB input should be logic-0, for subtraction sub input should be at logic-1 state.



Multiplexers & Demultiplexers

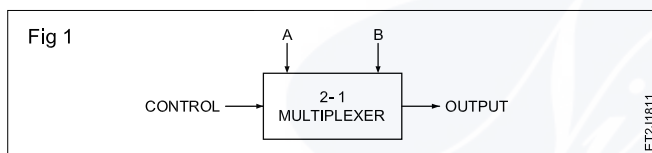
Objectives: At the end of this lesson you shall be able to

- explain the need of multiplexers and demultiplexers in digital circuits
- explain the function of a multiplexer with an example
- explain the relationship between number of input lines and required number of control lines
- list a few commercially available multiplexer and demultiplexer ICs
- explain the application of a multiplexer & demultiplexer in data transmission.

Many applications in digital logic requires circuit with multiple input and single output, single input and multiple outputs. The output of such circuits should however be uniquely determined by a set of control signals. Such circuits find immense application in computer and data transmission. Such circuits that have one or more input lines and give one or more output which are uniquely determined by the inputs are called *Combinational circuits*. Two of the most important combinational circuits are the *Multiplexers* and *Decoders*.

Multiplexers

A multiplexer having 2^n data inputs, one data output and an n -bit control input which selects one of the input and routes it to the output is shown in Fig 1.



In Fig 1, the multiplexer has two inputs ($2^n = 2^1=2$, hence $n=1$). It has 1-bit control signal (because, $n=1$) which selects A or B as the output as given in the Truth Table 1.

Truth Table

INPUTs		Control	Output
A	B		
1	0	0	1 (A --> output)
1	0	1	0 (B --> output)

The inverse of a Multiplexer is a Demultiplexer as shown in Fig-2. This has n input (in this case, $n=1$), $2n$ output (in this case, $2^1=2$ outputs) and n number of control signals (in this case $n=1$, hence control line=1). The single input is routed to one of the 2^n outputs, depending on the value of the n control lines. The truth table for the demultiplexer at Fig 2 is given in Table 2.

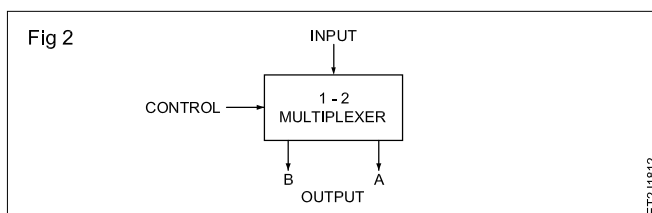


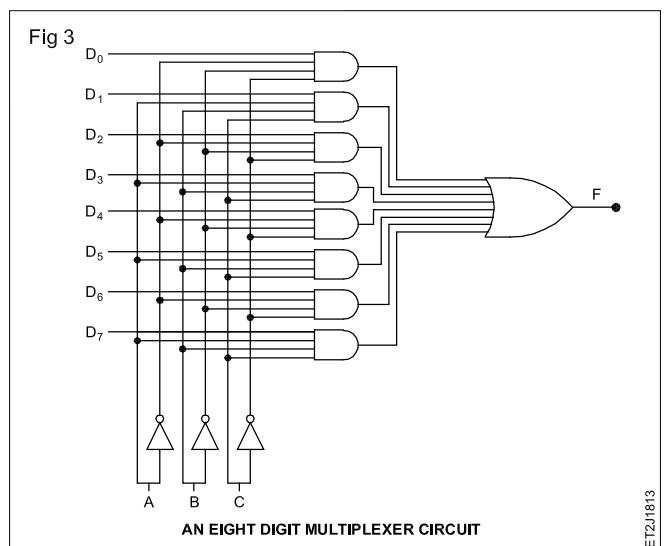
Table 2

INPUT	Control	Output
1	0	Input --> A (Therefore, A=1)
1	1	Input --> B (Therefore, B=1)

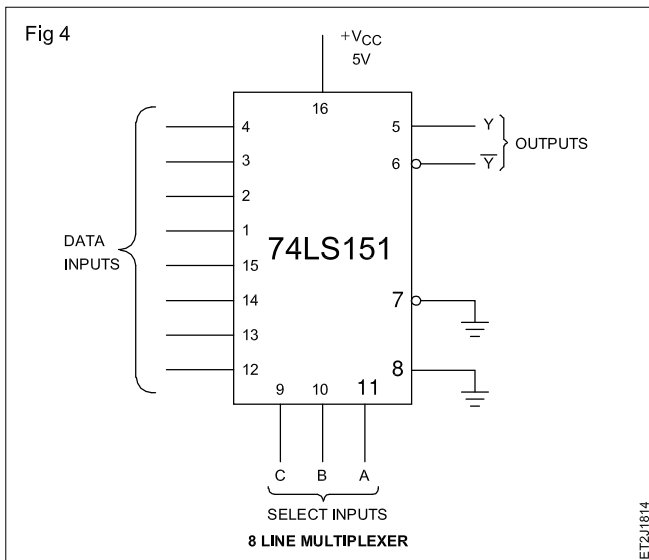
8-line Multiplexer

As discussed in earlier paragraphs, a multiplexer is a circuit with 2^n data inputs, one data output and n control inputs. The selected data is gated or routed to the output. Fig 3 shows the schematic of an eight-input or eight-line multiplexer.

As can be seen in Fig 3, the three control lines A, B and C encode a 3-bit number that specifies which of the eight input lines is gated to the OR gate and then to the output. Immaterial of what value is on the control lines, seven of the AND gate will always output 0, the other one may output 0 or 1 depending on the value of the selected input line. Each gate is enables by a different combination of the control inputs.



Such a eight-line multiplexer is available as a MSI chip. With 8 input lines, 3 control lines, one output, may be an additional compliment output line and power supply and ground lines is implemented as a 16 pin package. One such package is the 74LS151 , 8-line multiplexer IC shown in Fig 4.



Demultiplexer

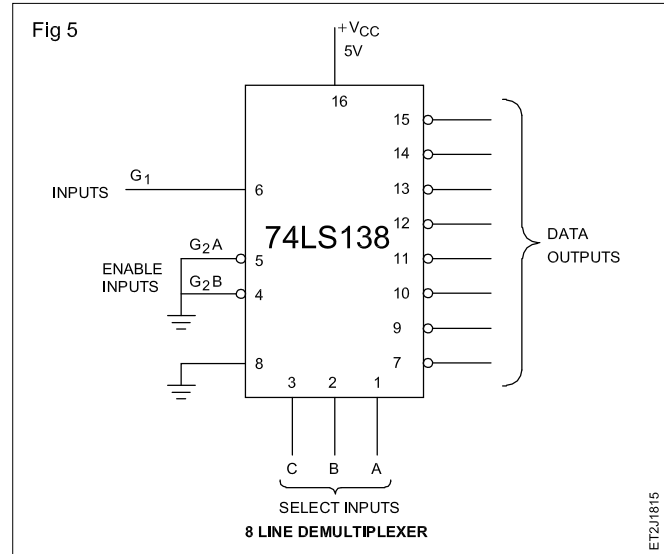
The inverse of a multiplexer is a demultiplexer. A demultiplexer routes its single input signal to one of 2^n outputs, depending on the values of the n control lines. For instance, if the binary value on the control signal is all zeros, the 0th output line is selected and if the binary value on the control lines is k , then, the k^{th} output line is selected for routing the input signal. Such demultiplexers are also available in IC package. One such IC is the 1line to 8 line demultiplexer 74LS138 as shown in Fig 5.

Application of Multiplexers and Demultiplexers

There are almost innumerable applications of multiplexers and demultiplexers. Just to list a few are in implementing a multiplexed display, parallel to serial data converter etc.,

The application of multiplexer and demultiplexer can be appreciated in data transmission as shown in Fig 6.

In Fig 6, the eight inputs could be eight signals coming from different transducers measuring eight different type of data (say, temperature, pressure,...) in a industrial environment. At the other end the output of the demultiplexer may be fed

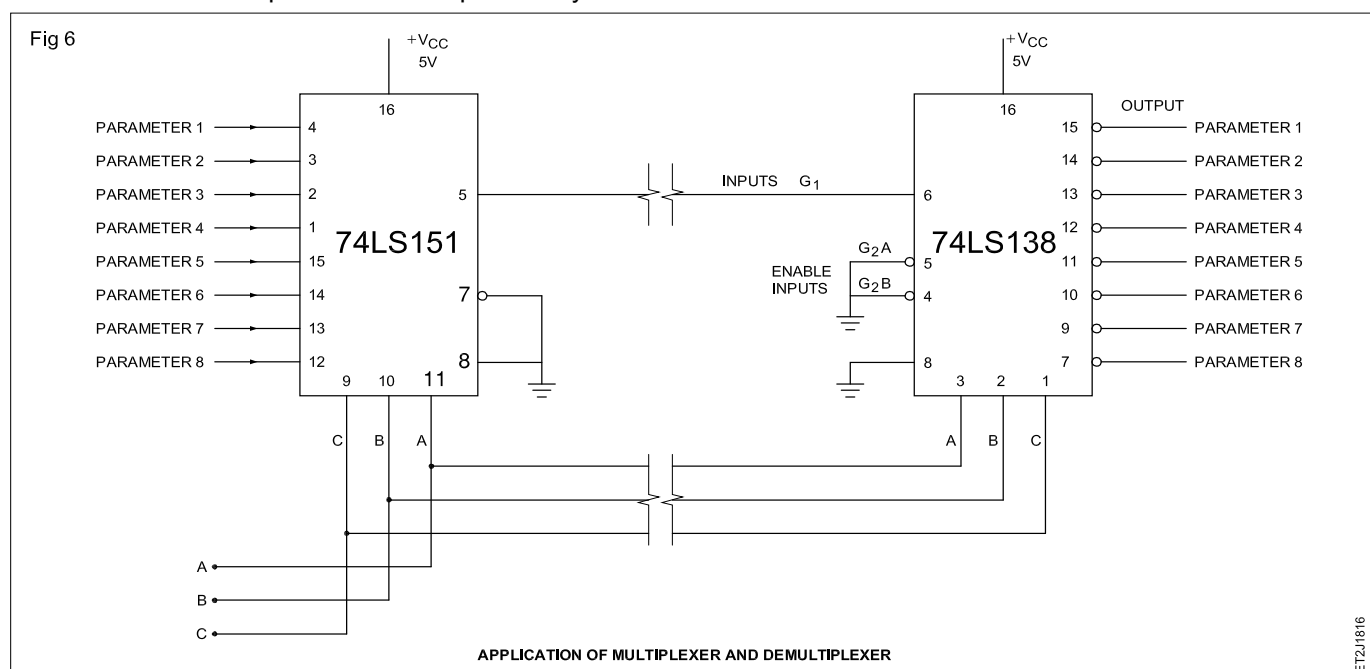


to eight different measuring instruments meant for measuring the individual parameters.

If the control lines of the multiplexer and demultiplexer are simultaneously fed with binary signals sequentially from 000 to 111, then each of the parameter of the input at any given time is communicated over the line to the demultiplexer which in-turn routes it to meter which is meant for displaying the value of the value of the parameter.

Observe from Fig 6, that only one transmission line is used for communicating all the eight parameters at different intervals of time. This is known as *Time division multiplexing*. Hence, multiplexers and demultiplexers are invariably used in such communication. The three control lines shown in Fig 6 could even be generated at sending and receiving station independently using one of the input line as the synchronizing input.

Discuss in detail the working of the schematic diagram shown in Fig 1 of exercise 19.19 taking different control signal situations.



Analog-to-digital converter

Objectives: At the end of this lesson you shall be able to

- explain the necessity of A to D converter
- explain successive-approximation method of A/D converter
- explain the characteristic of ADC0809 IC.

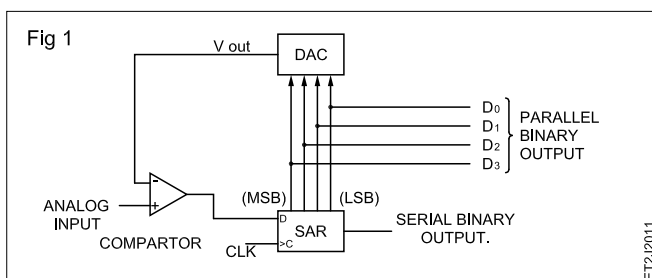
Analog to digital conversion is the process by which an analog signal is converted to an equivalent digital signal. This conversion is necessary when measured quantities must be in digital form for processing in a computer for display or storage. For example, an A/D converter is used to change the analog output signals from transducers measuring parameter like temperature, pressure, vibration etc into equivalent digital signals. The major factors that determine the quality performance of A/D converter are resolution, sampling rate, speed and linearity. The resolution is the smallest change in voltage that can be detected by the system and that can produce a change in the digital code. The speed of a A/D converter is determined by the time it takes to perform the conversion process. The sampling rate is the no. of times per second that the analog signal can be sampled and converted into a digital code. For proper A/D conversion, the minimum sampling rate must be atleast two times the highest frequency of the analog signal being sampled to satisfy the sampling theorem. The A/D requires a voltage reference in order to achieve absolute conversion accuracy. Same A/D converter ICs have internal voltage reference whereas others accept external voltage references. Analog signal can be converted to digital codes by many methods of which successive approximation and flash A/D conversion methods are most common. S/A - A/D conversion techniques is used commonly in medium to high speed data acquisition applies. It is one of the fastest A/D conversion techniques that requires a minimum amount of circuitry.

Successive-approximation A/D converter

Successive approximation is perhaps the most widely used method of A/D conversion. It has a much shorter conversion time, it also has fixed conversion time that is the same for any value of the analog input.

The Fig 1 shows basic block diagram of a 4 bit successive approximation ADC. It consists of a DAC, a successive-approximation register (SAR), and a comparator.

The basic operation is as follows:



The input bits of the DAC are enabled (made equal to logic-1) one at a time, starting with MSB, as each bit is enabled, the comparator produces an output that indicates whether the analog input voltage is greater or less than the output of the DAC for the corresponding I/p. If the DAC output is greater than the analog input, the comparator's output is low, causing the bit in the register to RESET. If the DAC output is less than the analog input the '1' bit is retained in the SAR register. The system does this with the MSB first, then the next most significant bit, then the next and so on. After all the bits of the DAC have been tried, the conversion cycle is complete.

Fig 2 illustrates the step-by-step conversion of a constant analog input voltage (5V in this case). Let us assume that the DAC has the following output characteristic: $V_{out} = 8V$ for the 2^3 bit (MSB), $V_{out} = 4V$ for the 2^2 bit, $V_{out} = 2V$ for the 2^1 bit and $V_{out} = 1V$ for the 2^0 bit (LSB).

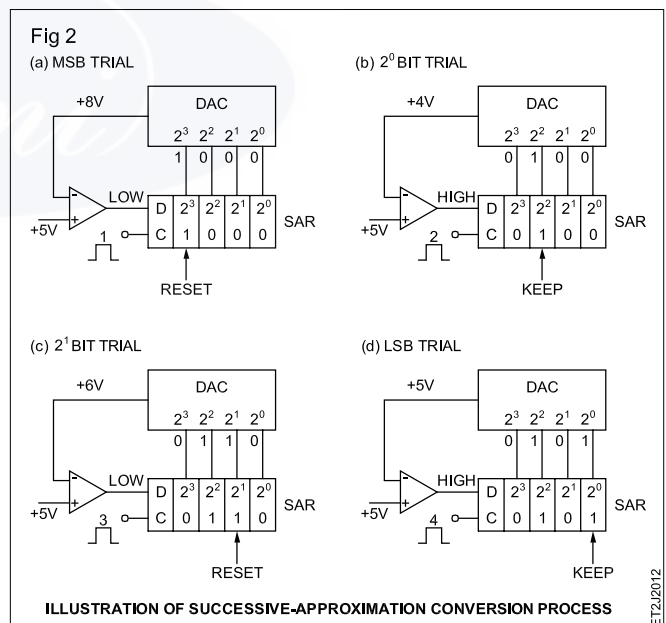


Fig 2a shows the first step in the conversion cycle with the MSB=1, the output of the DAC=8V. Since this is greater than the analog input of 5V, the output of the DAC is 8V. Since this is greater than the analog input of 5V, the output of the comparator is low, causing the MSB in the SAR to be reset to a logic-0.

Fig 2b shows the second step in the conversion cycle with the 2^2 bit equal to a logic-1. The output of the DAC is 4V, since this is less than the analog input of 5V, the output of the comparator switches to a HIGH, causing this bit to be retained in the SAR.

Fig 2c shows the third step in the conversion cycle with the 2^1 bit equal to a logic-1. The output of the DAC is 6 volts because there is a logic-1 on the 2^2 bit input and on the 2^1 bit input $4V+2V=6V$, since this is greater than the analog input of 5V, the output of the comparator. Switches to a LOW, causing this bit to be RESET to a logic-0.

Fig 2d shows the fourth and final step in the conversion cycle with the 2^0 bit equal to a logic-1. The output of the DAC is 5V because there is a logic-1 on the 2^2 bit input and on the 2^0 bit input $4V+1V=5V$.

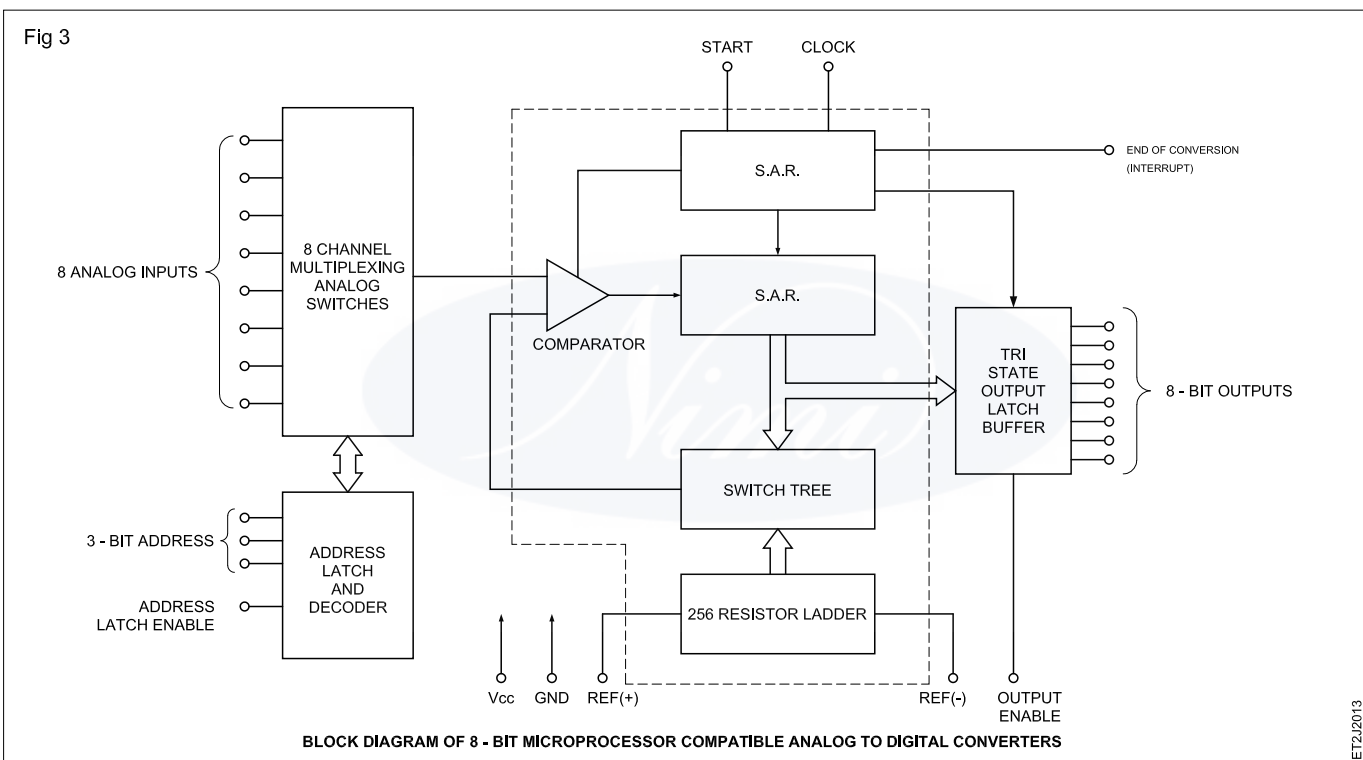
The four bits have all been tried, thus completing the conversion cycle. At this point the binary code in the register is 0101, which is the binary value of the analog input of 5V. Another conversion cycle now begins and the basic process is repeated. The SAR is cleared at the beginning of each cycle.

Analogue to digital converter ADC0808/0809 8 bit microprocessor compatible A/D converter

The ADC0808/0809 data acquisition device implement on a single chip most the elements of the standard data acquisition system. They contain an 8-bit A/D converter, 8 channel multiplexer with an address input latch and associated control logic. These device provide most of the logic to interface to a variety of microprocessors with the addition of a minimum number of parts.

Functional description

The ADC0808/0809 shown in the above Fig 3 can be functionally divided into 2 basic sub circuits. These two sub circuits are an analog multiplexer and an A/D converter. The multiplexer uses 8-standard CMOS analog switches to provide for upto 8 analog inputs, the switches are selectively turned on, depending on the data latched into 3-bit multiplexer address register.



The second function block, the successive approximation A/D converter transforms the analog output of the multiplexer to an 8 bit digital word, the output of the multiplexer goes to one of two comparator inputs. The other input is derived from a 256R resistor ladder, which is tapped by a MOSFET transistor switch tree. The converter control logic controls the switch tree, funneling a particular tap voltage to the comparator, based on the result of this comparison, the control logic and the successive approximation register (SAR) will decide whether the next tap to be selected should be higher or lower than the present tap on the resistor ladder, this algorithm is executed 8 times per conversion, once every 8 clock periods, yielding a total conversion time 64 clock period.

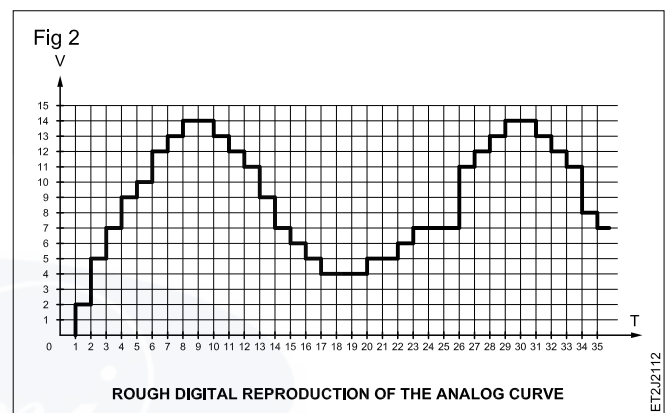
When the conversion cycle is complete the resulting data is loaded into the tri-state output latch can then be read by the host system any time before the end of the next conversion. The try state capability of the latch allows easy interface to bus oriented systems.

Objectives: At the end of this lesson you shall be able to

- differentiate between analog and digital system
- appreciate digital system over analog system
- explain different methods used in D to A conversion
- explain performance characteristic of D to A converter
- explain characteristics of DAC 0808IC.

An approximation of the analog function can be reconstructed from the sequence of digital numbers that has been generated, obviously, there will be some error in the reconstruction because only certain values are represented, Thirty six shown in the Fig 2 and not the continuous set of values.

In the case of a digital representation of the voltage using a 4-bit binary code, only sixteen values can be defined. More values between 0 and +15 can be represented by using more bits in the digital code. So an analog quantity can be represented to some degree of accuracy with a digital code that specifies discrete values within the range. This concept is illustrated in Fig 1.

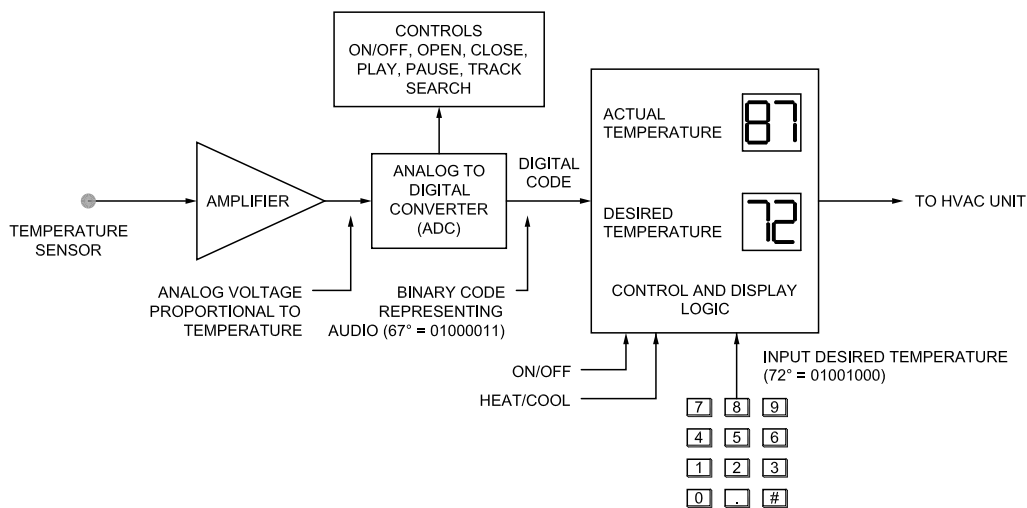


The following examples illustrate applications of these conversion processes.

The Fig 3 shows the simplified block diagram of an electronic thermostat, the room temperature sensor produces an analog voltage that is proportional to the temperature. This voltage is increased by the linear amplifier and applied to the analog-to-digital converter (ADC) where it is converted to a digital code and periodically sampled by the control logic.

345

Fig 3



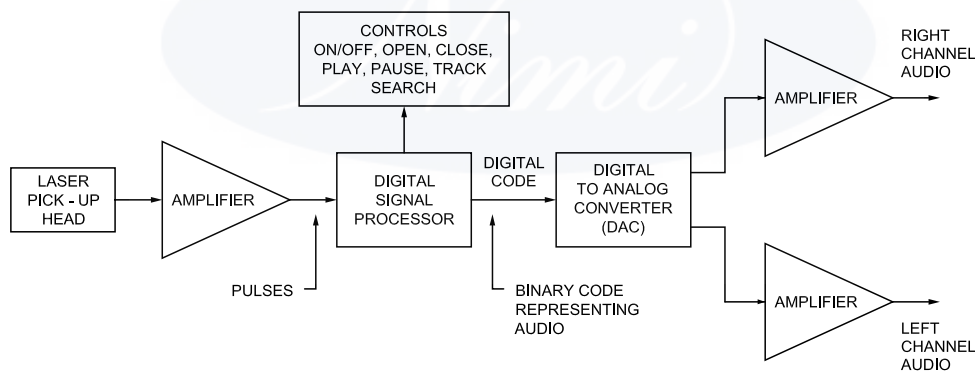
BASIC BLOCK DIAGRAM OF AN ELECTRONIC THERMOSTAT

ET2J2103

For Example: The room temperature is 67°F. A specific voltage value corresponding to this temperature appears on the ADC input and is converted to an 8-bit binary number 01000011. The control logic compares this binary number with a binary number representing the desired temperature (say 01001000 for 72°F). This desired value has been previously entered from the key pad and stored in a register. The comparison shows that the actual room temperature is less than the desired temperature. As a

result, the control logic instructs the control unit circuit to turn the furnace on. As the furnace runs, the control logic continues to monitor the actual temperature through the ADC. When the actual temperature equals or exceeds the desired temperature, the control logic turns the furnace off, both the actual temperature and the desired temperature are displayed.

Fig 4



BASIC BLOCK DIAGRAM OF A CD PLAYER

ET2J2114

CD player

The compact-disk (CD) player is an example of a system that uses a digital to analog converter (DAC). The block diagram is shown in Fig 4, an audio signal is digitally recorded on the CD in the form of pits that are sensed by the laser pick up and amplified. The amplified digital signal is converted to a sequence of binary codes that represent the originally recorded audio signal, the binary representation of the audio is converted to analogue form by the DAC, amplified, and sent to the speakers.

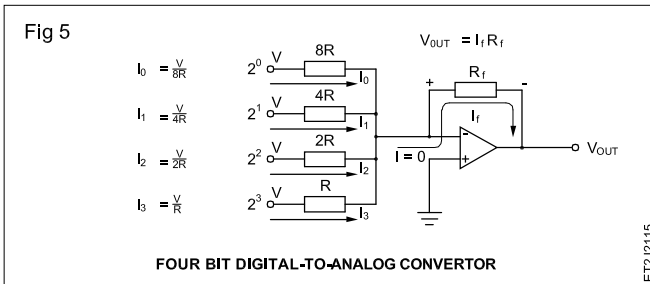
Digital-to-analog converter

Digital to analog converter is an important section in any digital system as indicated in the above two examples. OP-AMP is the most common element used in the D to A converter.

The basic problem in converting a digital signal into an equivalent analog signal is to change the 'n' digital voltage levels into one equivalent analog voltage. This can be most easily accomplished by designing a resistive network as shown in Fig 5, that will change each digital level into an equivalent binary weighted voltage or current.

The values of the input resistors network is chosen to be inversely proportional to the binary weights of the corresponding input bits, the lowest-value resistor (R) corresponds to the highest binary weighted input (2^n). The other resistors are multiples of R ($2R$, $4R$ and $8R$) and correspond to the binary weights, 2^{n-1} , 2^{n-2} , 2^{n-3} 2^0 respectively.

The I/P currents are also proportional to the binary weights, thus the o/p voltage is proportional to the sum of the binary weights because the sum of the input current is through R_f .



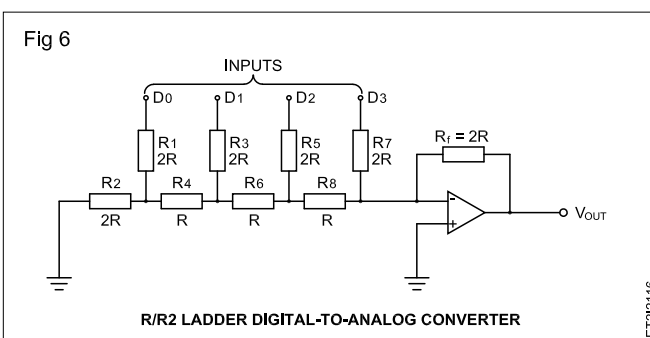
The Fig 5 shows a 4 bit DAC, each of the input resistors will either have current or have no current, depending on the input voltage level. If the input voltage is zero (binary 0), the current is also zero. If the input voltage is HIGH (binary 1), the amount of current depends on the input resistor value and is different for each input resistor.

Since there is practically no current into the op-amp inverting input, all of the input currents SUM together and go through R_f , since the inverting input is at 0V (Virtual ground), the drop across R_f is equal to the o/p voltage, so $V_{out} = I_f R_f$.

The main disadvantage of this type of DAC is the number of different resistor values. For example, An 8 bit converter requires eight resistors, ranging from some value of R to $128R$ in binary-weighted steps. This range of resistors requires tolerances of one part in 255 (less than 0.5%) to accurately convert the input, making this type of DAC very difficult to mass-produce.

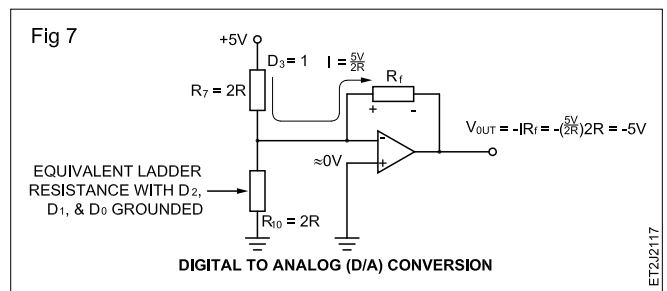
R/2R ladder digital-to-analog converter

The Fig 6 shows another method of DA conversion using R/2R ladder resistor network, for four bits. It overcomes one of the problems in the binary-weighted input DAC. In this type DAC only two values of resistors are required R and $2R$.



Assuming that the D_3 input is HIGH (+5V) and the others are LOW (ground 0V). This condition represents the binary 1000. A circuit analysis will show that this reduces to the equivalent form shown in Fig 7. Essentially no current goes through the $2R$ equivalent resistance because the inverting input is at virtual ground.

Thus all of the current ($I = 5V/2R$) through R_7 also goes through R_f , and the output voltage is -5V. The operational amplifier keeps the inverting (-) input near zero volts (~0V) because of -ve feedback. Therefore all current goes through R_f rather than into the inverting input.



When the D_2 input is at +5V and the others are at ground. This condition represents 0100. Thevenize the $R_2/2R$ ladder network looking from R_8 , results in a current through R_f of $I = 2.5V/2R$, which gives an output voltage of -2.5V. Keep in mind that there is no current into the Op-Amp inverting input and that there is no current through the equivalent resistance to ground because it has zero volts across it, due to the virtual ground.

When the input is 0010

When D_1 input is at +5V and the others are at ground, this condition represents 0010, again thevenizing the $R_2/2R$ ladder network looking from R_8 , results in a current through R_f of $I = 1.25V/2R$, which gives an output voltage of -1.25 volts.

When the input is 0001

When D_0 input is at +5V and the others are at ground, this condition represents 0001, again thevenizing the $R_2/2R$ ladder network looking from R_8 , results in a current through R_f of $I = 0.625V/2R$, which gives an output voltage of -0.625V.

Notice that each successively lower weighted input produces an o/p voltage that is halved, so that the output voltage is proportional to the binary weight of the input bits.

In equation form the o/p voltage is given by

$$V = \frac{D_0 2^0 + D_1 2^1 + D_2 2^2 + D_3 2^3 + \dots D_{n-1} 2^{n-1}}{2^n}$$

where $D_0, D_1, D_2, D_3, \dots, D_{n-1}$ are the digital input levels.

Performance characteristics of digital-to-analog converter

Resolution

The resolution of a DAC is the reciprocal of the number of discrete steps in the output. This of course, is dependent on the number of input bits.

For example: A 4-bit DAC has a resolution of one part in $2^4 - 1$ (one part in fifteen) expressed as a percentage, this is $(1/15) \times 100 = 6.67\%$. The total number of discrete steps equals $2^n - 1$, where n is the number of bits. Resolution can also be expressed as the number of bits that are converted.

Accuracy

Accuracy is a comparison of the actual output of a DAC with the expected output. It is expressed as a percentage of a full-scale, or maximum output voltage.

Example: If a converter has a full-scale output of 10V and the accuracy is $\pm 0.1\%$, then the maximum error for any output voltage is $(10V)(0.001) = 10\text{mV}$, ideally, the accuracy should be, at most $\pm 1/2$ of an LSB. For an 8-bit converter, 1 LSB is $1/256 = 0.0039$ (0.39% of full scale), the accuracy should be approximately $\pm 0.2\%$.

Linearity

A linear error is a deviation from the ideal straight-line output of a DAC. A special case is an offset error, which is the amount of output voltage when the input bits are all zeros.

Monotonicity

A DAC is monotonic if it doesn't take any reverse steps when it is sequenced over its entire range of input bits.

Settling time

Settling time is normally defined as the time it takes a DAC to settle within $\pm 1/2$ LSB of its final value when a change occurs in the input code.

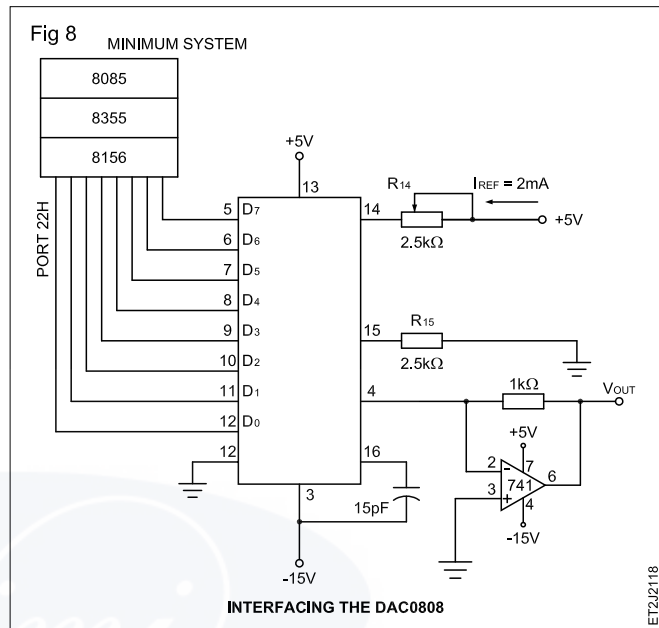
DAC 0808 Digital to Analog converter

The DAC 0808 series is an 8 bit monolithic digital-to-analog converter (DAC) featuring a full scale output current setting time of 150ns while dissipating only 33mw with $\pm 5\text{V}$ supplies. No reference current (I_{ref}) trimming is required for most applications since the full scale output current is typically ± 1 LSB of $255 I_{\text{ref}}/256$. Relative accuracies of better than $\pm 0.19\%$ assure 8-bit monotonicity and linearity, while zero level output current of less than $4\mu\text{A}$ provides 8-bit zero accuracy for $I_{\text{ref}} = 2\text{mA}$, the power supply currents of the DAC 0808 series are independent of bit codes and exhibits essentially constant device characteristics over the entire supply voltage range.

The DAC 0808 will interface directly with popular TTL, DTL or CMOS logic levels, and is a direct replacement for the MC1508/MC1408.

The Fig 8 shows the data bits of a DAC0808 connected to port 22H of minimum system. Pin 2 of the DAC0808 is grounded, and a 15pF compensating capacitor is between pins 16 and 3, +5V supply sets up a reference current for

the R-2R ladder, pin 14 is connected to a positive supply through a resistance (trimmer) R_{14} allows you to adjust this to 2mA and pin 15 is grounded through a resistance R_{15} is the same size as R_{14} ; this compensates for drift in the input stage of the converter. Notice that I_{out} drives the inverting input of an op-amp; therefore the final output ranges from 0 to +2V. Pin 1 is unused, pin 2 is chip ground, pin 3 (V_{EE}) is -15V, pin 4 is the ground return for the current out of the ladder; this pin usually connects to an op-amp, pin 5 to 12 are for the 8 bits of input data, Pin 13 (V_{CC}) is +5V. Finally a capacitor between pin 16 and pin 3 frequency-compensates the device.



Grounding and bypassing

D/A and A/D converter IC(s) require correct grounding and capacitive bypassing in order to operate according to performance specifications. The digital signals can severely impair analog signals to contact the electromagnetic interference caused by the digital signals, the analog and digital grounds should be kept separate and should have only one common point on the circuit board. Bypass capacitors are required at the power connections to the IC, the reference signal inputs and the analog inputs to minimize noise that is induced by the digital signals. Manufacturer specifies the recommended bypass capacitor locations and values in the data sheet.

Shift registers, Types, applications

Objectives: At the end of this lesson you shall be able to

- explain the basic functions of a shift register
- explain different configurations of shift register
- explain the function of IC7495 in different configurations.

INTRODUCTION TO SHIFT REGISTERS

A shift register is a very important digital building block. Registers are often used to momentarily store binary information appearing at the output of an encoding matrix. A register might be used to accept input data from an alphanumeric keyboard and then present this data at the input of a microprocessor chip. Similarly, shift registers are often used to momentarily store binary data at the output of a decoder. For instance, a register could be used to accept output data from a microprocessor chip and then present this data to the circuitry used to drive the display on a CRT screen. Thus registers form a very important link between the main digital system and the input-output channels.

A binary register also forms the basis for some very important arithmetic operations. For example, the operations of complementation, multiplication, and division are frequently implemented by means of a register. A shift register can also be connected to form a number of different types of counters. These counters offer some very distinct advantages.

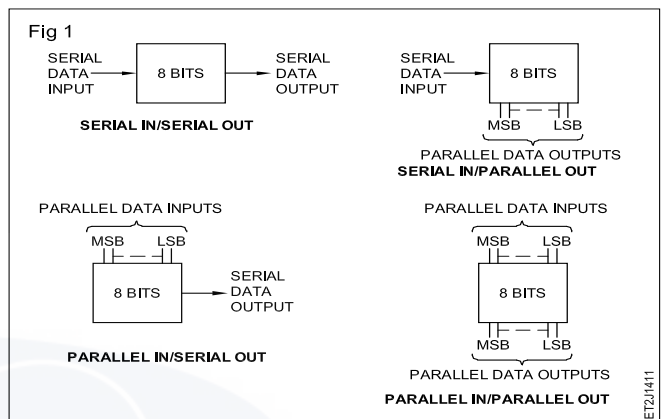
Types of registers

A register is simply a group of flip-flops that can be used to store a binary number. There must be one flip-flop for each bit in the binary number. For instance, a register used to store an 8-bit binary number must have eight flip-flops. Naturally the flip-flops must be connected such that the binary number can be entered (shifted) into the register and possibly shifted out. A group of flip-flops connected to provide either or both of these functions is called a shift register.

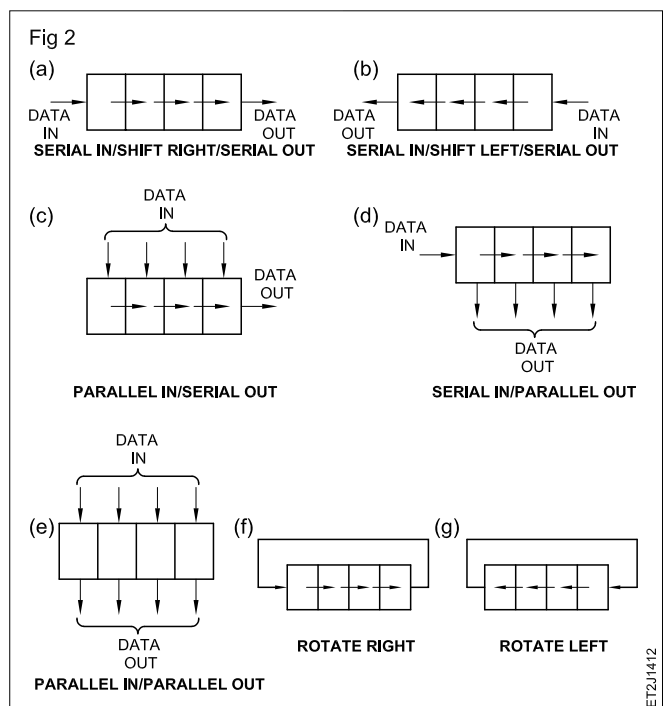
The bits in a binary number (let's call them the data) can be moved from one place to another in either of two ways. The first method involves shifting the data 1 bit at a time in a serial fashion, beginning with either the MSB or the LSB. This technique is referred to as serial shifting. The second method involves shifting all the data bits simultaneously and is referred to as parallel shifting.

There are two ways to shift data into a register (serial or parallel) and similarly two ways to shift the data out of the register. This leads to the construction of four basic register types as shown in Fig 1 -serial in - serial out, serial in - parallel out, parallel in - serial out, and parallel in - parallel out. All of these configurations are commercially available as TTL MSI/LSI circuits. For instance:

Serial in - serial out - 54/74L91, 8 bits
 Serial in - parallel out - 54/74164, 8 bits
 Parallel in - serial out - 54/75165, 8 bits
 Parallel in - parallel out - 54/74194, 4 bits
 Parallel in - parallel out - 54/74198, 8 bits



The shifting capability of a register permits the movement of data from stage to stage within the register or into or out of the register upon application of clock pulses. Fig 2a to 2g illustrate the types of data movement in shift registers. The block represents any arbitrary 4-bit register and the arrow indicate the direction of data movement.



Serial in - serial out operation

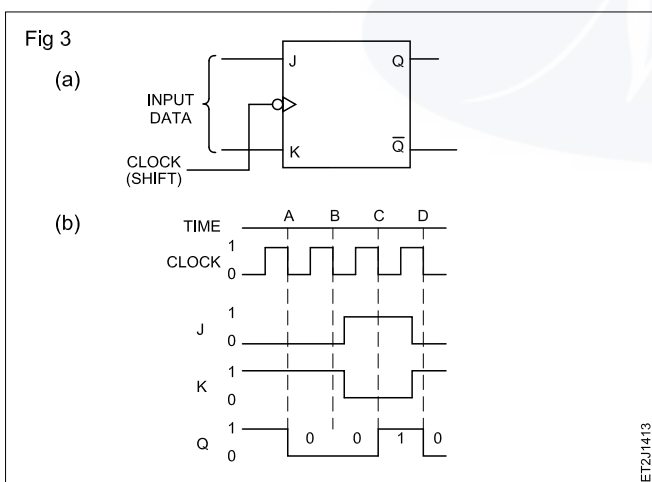
The flip-flops used to construct registers are usually either JK or D types. So let's begin by summarizing the operation of JK flip-flop.

For a JK flip-flop, the data bit to be shifted into the flip-flop must be present at the J and K inputs when the clock transitions (low or high). Since the data bit is either a 1 or a 0, there are two cases:

- To shift a 0 into the flip-flop, $J=0$ and $K=1$.
- To shift a 1 into the flip-flop, $J=1$ and $K=0$.

The important point to note is that the J and K inputs must be controlled to provide the correct input data. The J and K logic levels may be changing while the clock is high (or low), but they must be steady from just before until just after the clock transition (remember, setup time and hold time). For our discussion we shall use JK master-slave flip-flops having clock inputs that are sensitive to negative clock transitions. Incidentally, this negative transition of the clock is frequently referred to as a shift pulse.

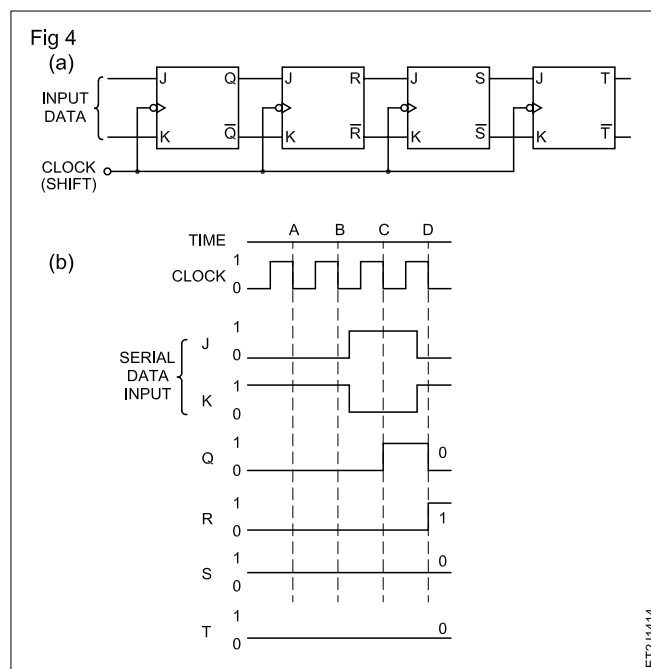
The waveforms in Fig 3 illustrate these ideas. At time A, Q is reset low (a 0 is shifted into the flip-flop). At time B, Q does not change since the flip-flop had a 0 in it and another 0 is shifted in. At time C, the flip-flop is set (a 1 is shifted into it). At time D, another 0 is shifted into the flip-flop. In essence, we have shifted 4 data bits into this flip-flop in a time sequence: a 0 at time A, another 0 at time B, a 1 at time C, and a 0 at time D.



Now, consider adding three more flip-flops connected as shown in Fig 4. Let's begin with all the flip-flops reset and then apply the exact same input signals to flip-flop Q as we did in Fig 3. Here's what happens:

At time A: All the flip-flops are reset, so all J inputs are low and all K inputs are high. Then T is reset (the 0 in S is shifted into T). Similarly, the 0 in R is shifted into S, the 0 in Q is shifted into R, and the 0 at the data input is shifted into Q. The flip-flop outputs just after time A are QRST = 0000.

At time B: The flip-flops all contain 0s. Thus the 0 in S is shifted into T, the 0 in R shifts into S, the 0 in Q shifts into R, and the 0 at the data input shifted into Q. The flip-flop outputs are QRST = 0000.



At time C: The flip-flops still all contain 0s. The 0 in S shifts into T, the 0 in R shifts into S, and the 0 in Q shifts into R, but a 1 at the data input now shifts into Q. The flip-flop outputs are QRST = 1000.

At time D: The 0 in S shifts into T, the 0 in R shifts into S, the 1 in Q shifts into R (the J input to R is high and the K input is low), and the 0 at the data input shifts into Q. The flip-flop outputs are QRST = 0100.

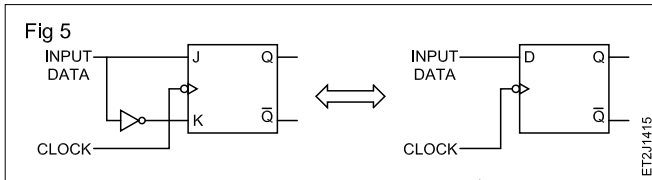
To summarize, we have shifted 4 data bits in a serial fashion into four flip-flops. These 4 data bits could represent a 4-bit binary number 0100, assuming that we began shifting with the LSB first. Notice that the LSB is in T and the MSB is in Q. These four flip-flops could be defined as a 4-bit shift register; thus this is the technique used to construct a serial-input shift register.

The serial data input for the register shown in Fig 4 requires two input signals J and K. But look carefully at the waveforms. Clearly, $K = J$, or $J=K$. In other words, one signal is always the complement of the other. If we were to connect an inverter between J and K on flip-flop Q with the input at J, therefore, we would need to have only one data input signal - the one required for J. But this is precisely a D-type flip-flop as shown in Fig 5. Remember the rules for a type D flip-flop; on the negative clock transition, the data present at the D input (either a 1 or a 0) will shift into the flip-flop.

Thus the 4-bit serial input shift register shown in Fig 5 can be constructed by replacing the JK flip-flops with type D flip-flops.

Serial in parallel out (SIPO)

Data is entered serially into this type of register, and data bits are taken out of the register parallel from the output of each stage. Once the data bits are stored, each bit appears on its respective output line, and all bits are available simultaneously, rather than on a bit-by-bit basis as with the serial output.



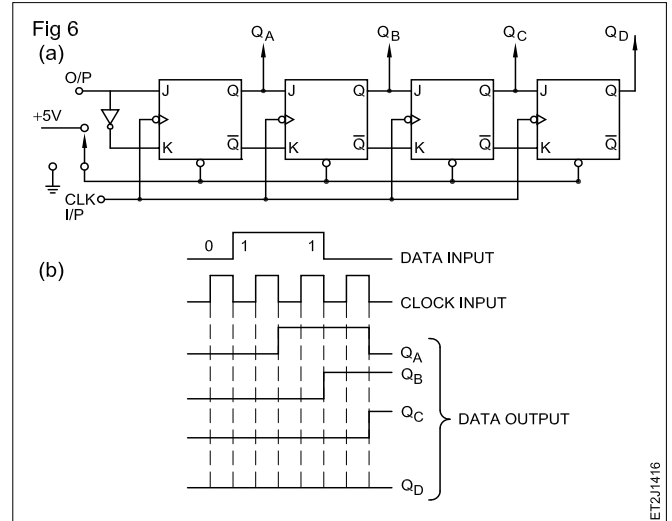
The Fig 6a shows the SIPO shift register using J.K flip-flops. All the inputs are tied to the complement of J-inputs. The clock, preset and clear inputs are -ve edge triggered.

Initially all the flip-flops are cleared to logic-0 state by applying logic-0 pulse to the clear inputs.

The sequence of shifting logic information is shown in Fig 6b which gives waveform diagrams of data input, data output and clock input.

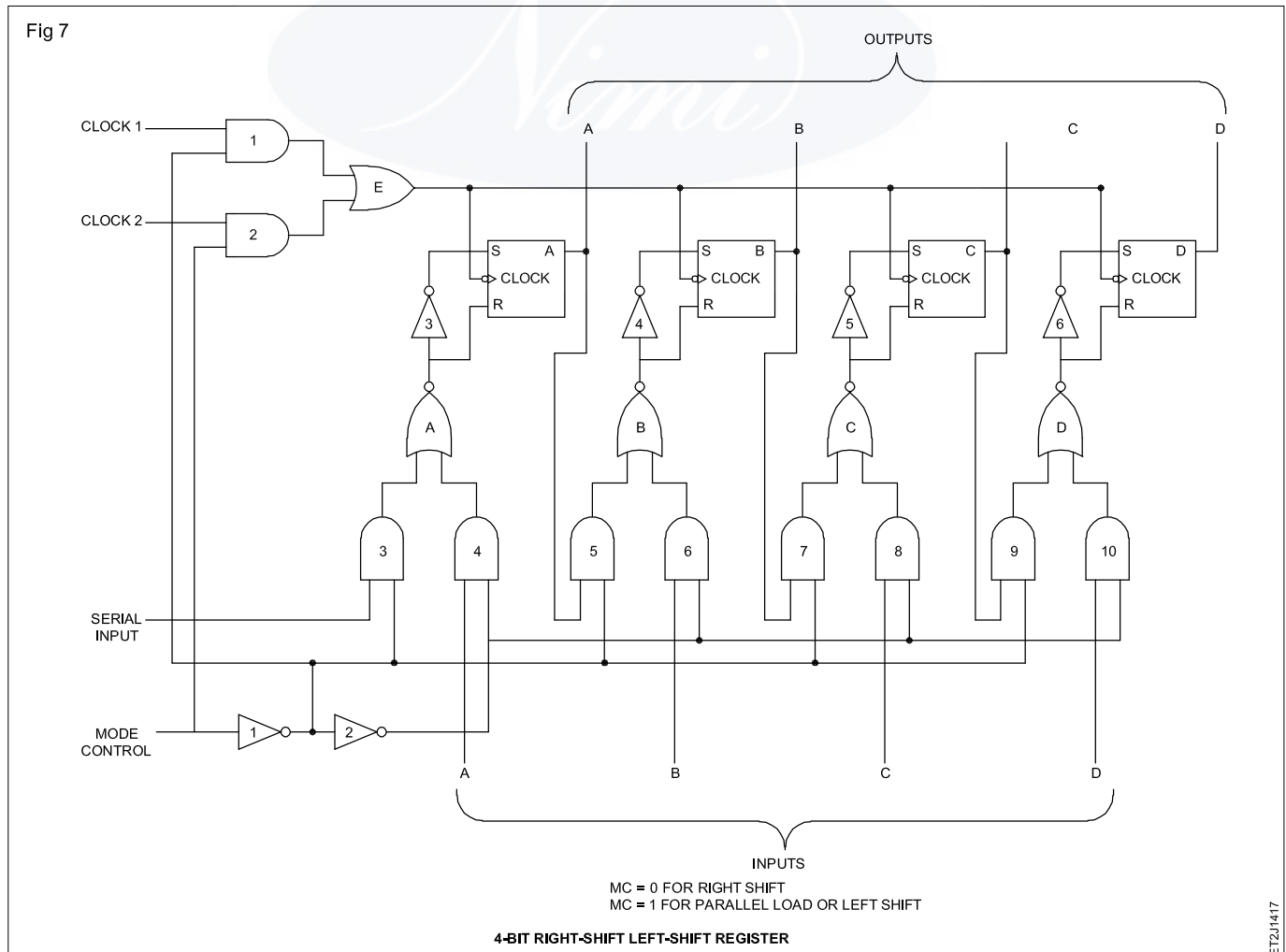
4-bit right-shift left-shift register IC7495

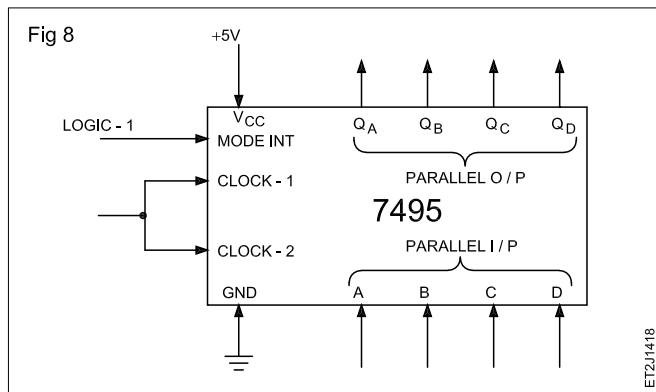
The Fig 7 shows the internal logic diagram of 4 bit shift register with parallel load capability, and with all flip-flop outputs available. This makes it possible to perform the right shift or left shift operation under control of the mode control input. For greater flexibility, the mode control selects clock-1 for the right shift mode and clock 2 for the parallel load (left shift) mode. The clock 1 and clock 2 inputs are tied together if only one clock source is required. Data transfer occurs once the -ve going edge of the clock pulse occurs.



Parallel in parallel out (PIPO)

The Fig 8 shows the pins to be used for parallel in parallel out operation in IC7495. For this operation mode control should be kept at logic high. Clock 2 should be selected for applying clock pulse, data inputs are to be connected to the A, B, C & D parallel inputs, data output can be taken simultaneously from QA, QB, QC and QD.





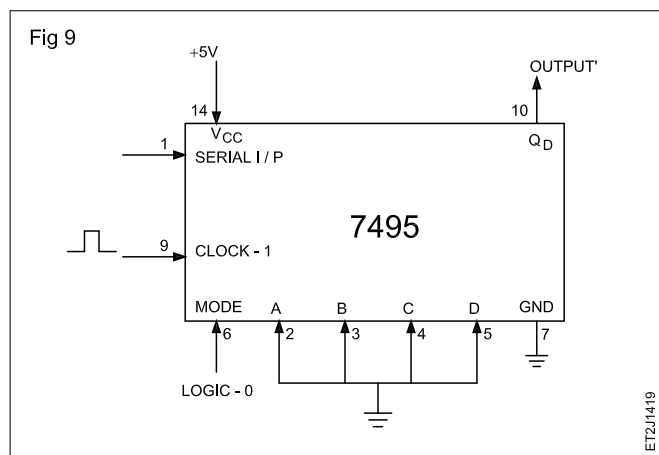
When mode control is at logic high, inverter 1 output will be at logic-0. Hence AND gates 1,3,5,7 and 9 are disabled because one of the inputs of those gates will be at logic-0. Therefore clock 1 and serial inputs will be disconnected from the flip-flops. At the same time AND gates 2,4,6,8 and 10 are enabled because, both the AND gate inputs of enabled gates will have high inputs simultaneously, when clock is high. Hence only clock 2 and parallel inputs are routed to the flip-flop inputs. Falling edge of the clock pulse, transmits the data information from parallel inputs to parallel outputs, simultaneously.

This type of register requires very less time (i.e one clock pulse) for transfer 1 set of parallel data information.

Serial in serial out (SISO)

The Fig 9 shows the pins to be used for SISO operation in IC 7495. When the mode control input is at logic-0 condition inverter 1 (refer functional diagram of 7495 for inter blocks) output will be at logic-1 and that enables AND gates 3,5,7,9 and 1. AND gate-1 selects clock-1 status for triggering flip-flops through OR gate. AND gates 3,5,7,9 selects serial input data. At the same time inverter-2 output will be at logic-0 that disables AND gates 4,6,8,10 thus disconnecting all parallel inputs A,B,C and D reaching flip-flop. 4 clock pulses are required to transfer data from input to output of the shift register.

If the outputs are taken from all the outputs Q_A , Q_B , Q_C and Q_D , then this register works like serial in parallel out (shift right) register.

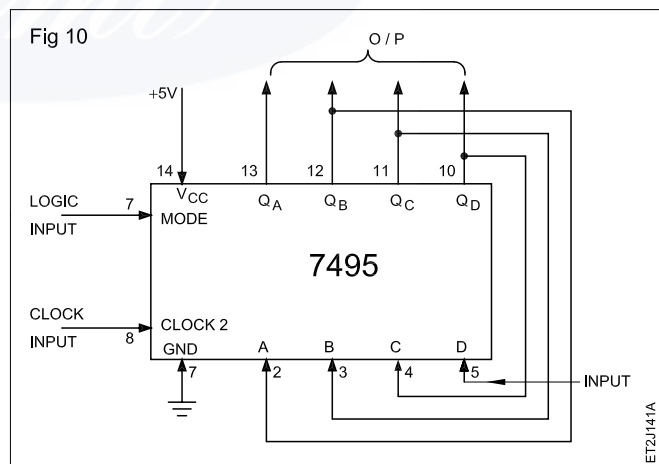


Serial in parallel out (shift left)

The Fig 10 shows external wiring diagram for serial in parallel out shift left register. For shift left operation mode control signal level should be logic-1.

Data input is applied to the parallel input-D remaining parallel inputs A, B and C receives signals from the outputs Q_B , Q_C and Q_D respectively as shown in Fig 10.

When mode control signal is at logic-1 AND gate 2 is enabled and AND gate 1 is disabled. Hence all the flip-flops get triggering pulse from clock 2. At the same time AND gates 3,5,7 and 9 are disabled and AND gates 4,6,8 and 10 are enabled, therefore parallel inputs A,B,C and D routed to the flip-flops inputs, and serial input is disabled from the flip-flops. The data input given at the D input is shifted left as the clock pulse progress. To move data from Q_D to Q_A , 4 clock pulses are required.



Gears, Belts, Stepper Motor

Objectives : At the end of this lesson you shall be able to

- define types of Gears
- define types of Belts
- define types of Stepper Motor.

Gears:

Gears are used in most of the mechanical devices. They do several important jobs, importantly, they provide speed increase or speed reduction in motorized equipment. This is important because, often, a small motor spinning very fast can provide enough power for a device, but not enough torque. For instance, in a Laser Printer has gear reduction because it needs lots of torque to turn the Drum in the Toner cartridge, but the motor only produces a small amount of torque at a high speed. With a gear reduction, the output speed can be reduced while the torque is increased.

Types of Gears

- 1 Spur Gears
- 2 Helical Gears
- 3 Herringbone Gears
- 4 Bevel / Miter Gears
- 5 Worm Gears
- 6 Internal Gears
- 7 Racks
- 8 Face Gears
- 9 Involute Splines
- 10 Straight Sided Splines
- 11 Sprockets

Spur Gears

Spur gears are by far the most common type of gear and with the exceptions of the "cog" the type of gear that has been around the longest. Spur gears have teeth that run perpendicular to the face of the gear.

Fig 1



Helical Gears

Helical gears are very similar to spur gears except the teeth are not perpendicular to the face. The teeth are at an angle to the face giving helical gears more tooth contact in the same area. Helical gears can also be used on non-parallel shafts to transmit motion. Helical gears tend to run quieter and smoother than spur gears due to the increased number of teeth in constant contact at any one period of time.

Fig 2



Herringbone Gears

Herringbone gears resemble two helical gears that have been placed side by side. They are often referred to as "double helicals". One benefit of herringbone gears is that it helps to avoid issues related to side thrust created with the use of helical gears.

Fig 3



Bevel / Miter Gears

Bevel gears are used mostly in situations that require power to be transmitted at right angles (or applications that are not parallel). Bevel gears can have different angles of application but tend to be 90°.

Fig 4



Worm Gears

Worm gears are used to transmit power at 90° and where high reductions are required. The worm resembles a thread that rides in concaved or helical teeth.

Fig 5



Internal Gears

Internal gears typically resemble inverted spur gears but are occasionally cut as helical gears.

Fig 6



Racks

A rack is basically a straight gear used to transmit power and motion in a linear movement.

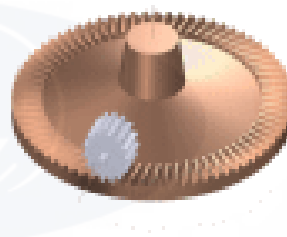
Fig 7



Face Gears

Face gears transmit power at (usually) right angles in a circular motion. Face gears are not very common in industrial application.

Fig 8



Involute Splines

Splined shafts and hubs are usually used as connectors in many different types of applications. One of the most common applications is to connect motors to gear reducers. They may also be used in transmissions. Involute splines resemble spur gears, but tend to have different pressure angles.

Fig 9



Straight Sided Splines

Straight sided splines often serve the same function as involute splines but have "straight sided" teeth instead of involute teeth.



Sprockets

Sprockets are used to run chains or belts. They are typically used in conveyor systems.



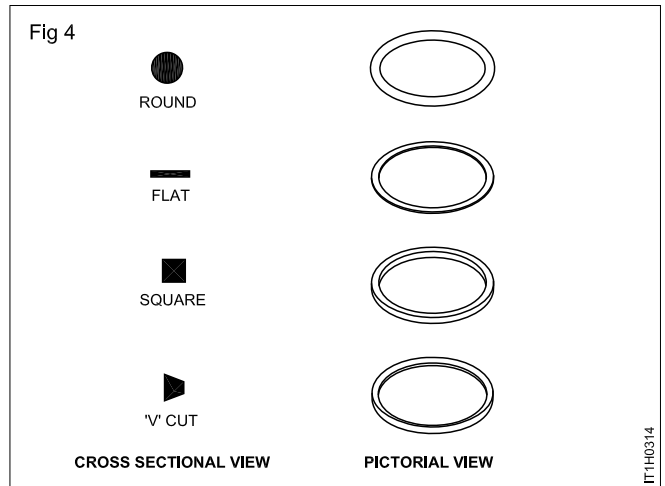
Belts

It is the main driving mechanism of all rotating parts. Equipments may have one or more belts to drive various rotating parts. Belt drives are widely used in power transmission. They are subject to rotary and push-pull motions with varying dynamic characteristics. Belts are friction drives, which mean they depend on friction between the belt and pulley/sheave to transmit power. Since it is the belt which transfers the rotation to all other rotating parts, the quality of belt used should be very good. Rubber is the main material used for making the belt. A belt made of low quality rubber may become loose very soon resulting in slipping. The sides of the belt may get worn out due to its constant rubbing with other parts. Due to this, the belt tension becomes weak with use. This will make the belt slip away from the pulley or it will create irregular speed in the mechanism. Also, a low quality belt may become sticky with time and temperature.

Selection of belts

Different types of belts are available in the market in various sizes as shown in Fig 12. Belts are generally available in four types. There are four types of belts available, they are:

- 1 Square type
- 2 Round type
- 3 V out type



- 4 Flat type
- 5 Toothed belt/timing belt

While replacing belt in any mechanism the following points should be considered;

- 1 Always use original size of belt.
- 2 DO NOT compromise on the quality of the belt.
- 3 DONOT use a different type of belt. (Donot use a V - type or Square type in place of a flat belt)
- 4 While changing the belt be sure about the pulleys through which it should be routed.
- 5 Donot elongate and use a belt.

A General information on stepper motors

Stepper motor

Stepper motors are special versions of the synchronous machine, in which the rotor is a permanent magnet, while the stator consists of a coil package. In contrast to synchronous motors, stepper motors have a large number of pole pairs. Motor operation requires a control unit, which energizes the individual motor windings based on a certain pulse sequence. A stepper motor has a tendency to mechanical oscillation. Above its load limit it loses dynamic characteristics and may lose individual steps. Under high load the shaft may even stop. Safe positioning is therefore only guaranteed within the performance limits. If the motor is operated within its rated limits, positioning without feedback of the rotor position can be achieved by linking individual steps. The operating mode (open loop control) and the durability of the stepper motor enable it to be used as a positioning drive in price-sensitive applications.

1 Basic function principles of a stepper motor

Like most electric motors, a stepper motor consists of a stator (fixed external winding) and a rotor (rotating shaft with magnets). The rotation of the motor shaft (rotor) is generated by rapid energizing the electromagnetic field of the stator, causing the shaft to turn by the step angle α . In a minimum control configuration, the stepper motor is moved from pole to pole, or from step to step.

A full turn of the motor shaft is therefore made up of individual steps. Energising of the motor windings results in a magnetic field in the motor from north to south (or south to north if the power supply has negative polarity and the winding is arranged accordingly). The movable stator with its permanent magnets aligns itself according to the direction of the external magnetic field of the stator.

Basic stepper motor types

Stepper motors come in numerous varieties, which tend to be based on three basic types:

- 1 Variable reluctance (stepper) motor (VR)
- 2 Permanent magnet stepper motor, claw - poled PM motor, Tin -Can - Motor (PM)
- 3 Hybrid stepper motor (HY)

Variable reluctance step per motor (Fig. 1, left)

The rotor consists of a magnetically soft material with different tooth pitch compared with the pole pitch of the stator. When a phase element is excited the rotor assumes a position in which the magnetic resistance (reluctance) for the excited magnetic circuit is at a minimum. Movement of the rotor causes a torque that returns the rotor to its original position. The stator requires at least two phase windings in order to achieve a change in the direction of

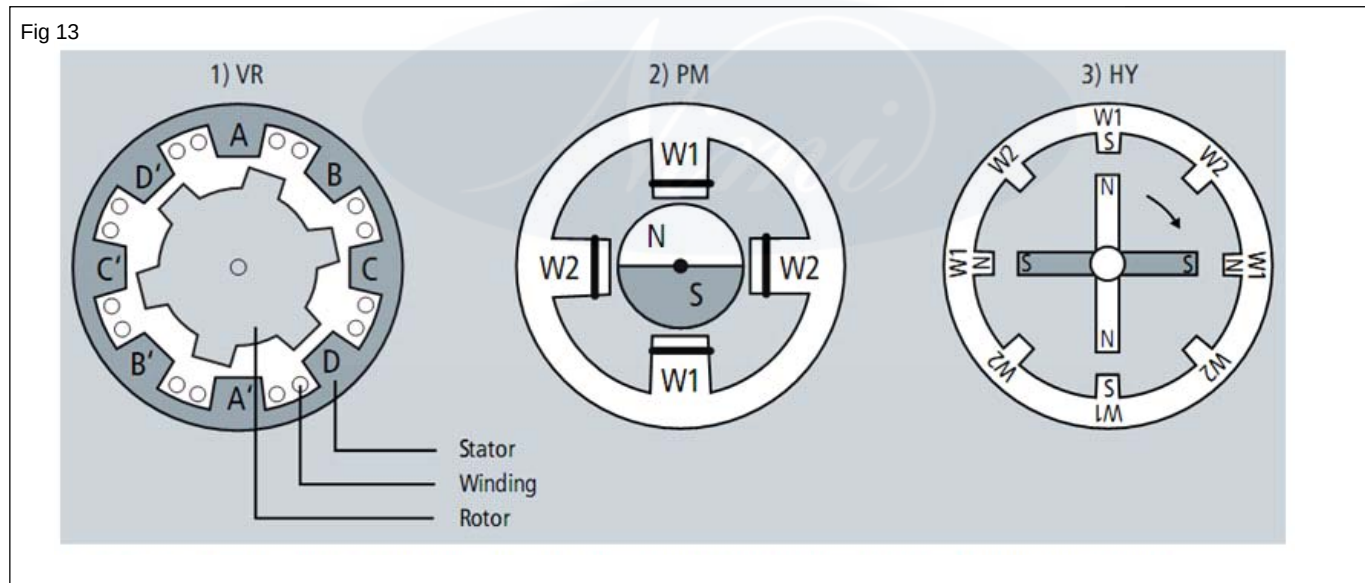
rotation. Due to the different tooth pitch of the pole and the rotor, the direction of the rotation field is opposite to the direction of rotation of the rotor. In de-energised state the motor has no holding torque.

Permanent magnet stepper motor /claw - pole stepper motor (Fig. 1, centre)

A permanent magnet stepper motor consists of a stator with individually controllable with magnet. The rotor may be configured as a cylindrical ferrite rod with multi-pole magnetisation along its circumference. The permanent-magnetic rotor always aligns itself with the right polarity relative to the stator winding. Permanent magnet stepper motors have a holding torque to which a non-energised motor can be subjected without causing a continuous rotation.

Hybrid stepper motor (Fig. 13, Right)

Combination of VR (Small step angles) and PM (high torque, holding torque). The rotor consists of a permanent magnet arranged in axial direction and positioned between two magnetically soft toothed discs. The toothed discs are offset by half a tooth pitch. Due to the arrangement of the permanent magnets in the rotor one of the toothed discs forms the north pole, the other the south pole of the rotor. The rotor teeth align themselves with the stator teeth, depending on which stator phase is current-carrying.



Electrical / Electronic & Mechanical Sensors

Objectives : At the end of this lesson you shall be able to

- types of electrical / electronic sensors
- types of mechanical sensors
- explain types of sensors.

Sensors

A sensor is a device that measures a particular characteristic of an object or system. Some sensors are purely mechanical, but most sensors are electronic, returning a voltage signal that can be converted into a useful engineering unit. Sensors take advantage of the mechanical or electrical response of its components to relate the response to a relevant quantity. Engineers use sensors in test and monitoring applications, but homeowners interact with sensors every day. Automobiles are filled with sensors, from the engine to the airbag.

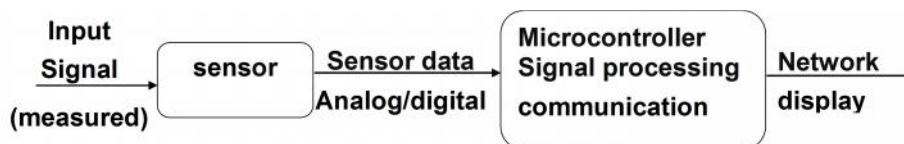
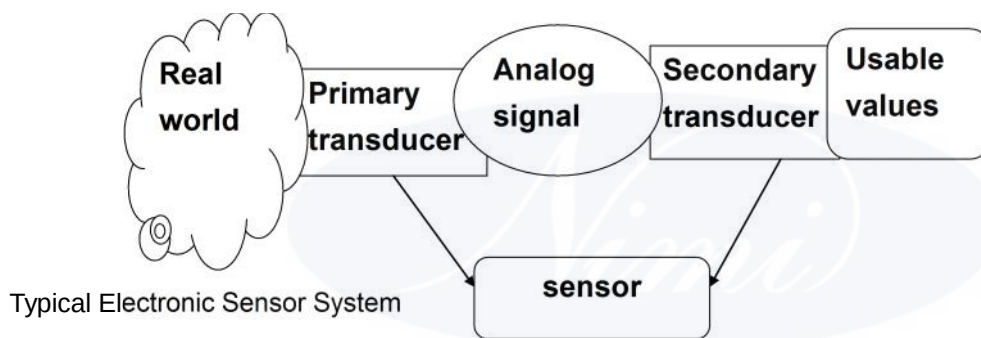
Electrical / Electronic Sensors

convert desired parameter into electrically measurable signal

General Electronic Sensor

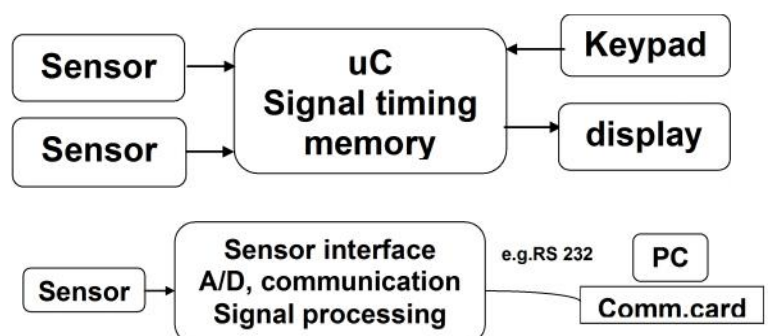
primary transducer: changes "real world" parameter into electrical signal

secondary transducer : converts electrical signal into analog or digital values

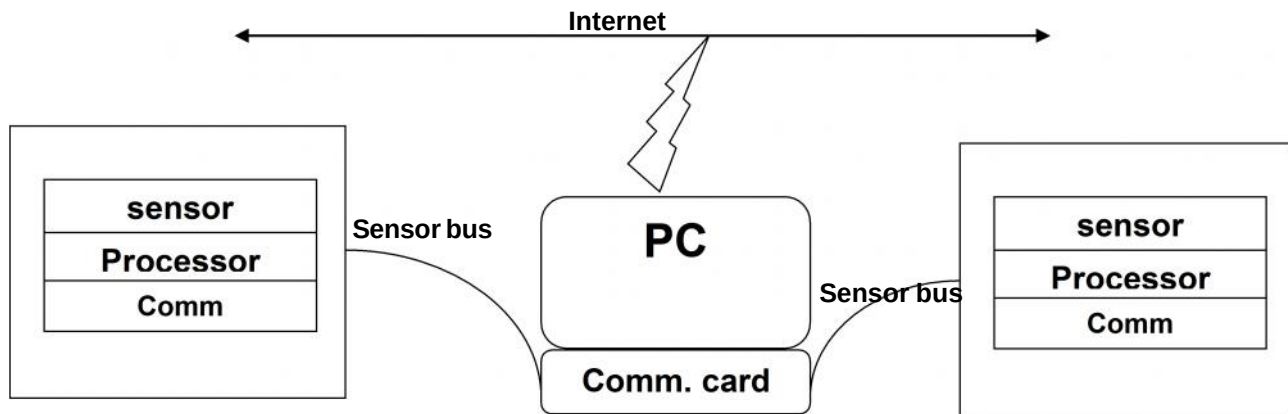


Example Electronic Sensor Systems

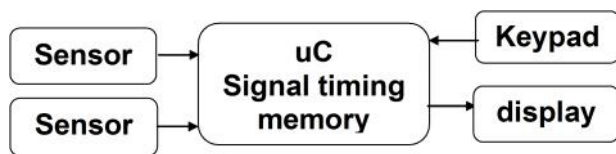
- Components vary with application
 - digital sensor within an instrument
- microcontroller
 - signal timing
 - data storage
 - analog sensor analyzed by a PC



Multiple sensors displayed over internet



Connecting sensors to microcontrollers



- Analog
- many microcontrollers have a built - in A/D
 - 8-bit to 12-bit common
 - many have multi-channel A/D inputs
- Digital
- Serial I/O
 - use serial I/O part, store in memory to analyze
 - synchronous (with clock)
- Must match byte format, stop/start bits, parity check, etc
 - asynchronous (no clock) : more common for comm. than data
- Must match baud rate and bit width, transmission protocol, etc
- **Frequency encoded**
 - use timing port, measure pulse width or pulse frequency
- Interface through lab view or user-generated code
 - Communication links common for sensors
- asynchronous serial comm
 - Universal asynchronous receive and transmit (UART)
 - 1 receive line + 1 transmit line. nodes must match baud rate & protocol
 - RS232 serial port on PCs uses UART format (but at +/-12v)
 - Can buy a chip to convert from UART to RS232
- **Synchronous serial comm**
 - Serial peripheral interface (SPI)
 - 1 clock + 1 bidirectional data + 1 chip select / enable
 - I²C = Inter integrated circuit bus
 - Designed by Philips for comm. inside TVS, used in several commercial sensor systems
- IEC 61451 : Sensor comm. standard
 - Several different sensor comm. protocols for different applications

Connecting smart sensors to PC/Network

- “smart sensor”= sensor with built-in signal processing & communication
- e.g., combining a “dumb sensor” and a microcontroller
 - Data acquisition cards (DAQ)
- PC card with analog and digital I/O

Mechanical Sensors

Mechanical sensors measure the change in a mechanical property of an object or system. The primary mechanical sensor is the strain gauge, which forms the basis of several different types of mechanical sensors. A strain gauge is a variable resistor that measures the amount of deformation that a part endures when it is affected by a force. Strain gauges form the basis of load cells, humidity sensors and pressure transducers. Another common mechanical sensor is the potentiometer, which measures angular or linear displacement.

Mechanical sensors are used for positioning and limit switch tasks on machine tools and presses, flexible production centres, robots, assembly and conveying equipment, and in machine and plant construction. For decades, they have proven their worth as the traditional strongmen of automation.

- Reliable and robust: even in harsh environments
- Flawless function in the face of vibrations, shocks, rapid temperature fluctuations, aggressive cooling lubricants, and heavy chip accumulation

Piezoelectric sensors

Piezoelectricity is the ability of certain crystals to produce a voltage when subjected to mechanical stress. The word is derived from the Greek *piezein*, which means to squeeze or press. The effect is reversible; piezoelectric crystals, subject to an externally applied voltage, can change shape by a small amount. The effect is of the order of nanometres, but nevertheless finds useful applications such as the production and detection of sound, generation of high voltages, electronic frequency generation, and ultrafine focusing of optical assemblies.

In a piezoelectric crystal, the positive and negative electrical charges are separated, but symmetrically distributed, so that the crystal overall is electrically neutral. When a stress is applied, this symmetry is disturbed, and the charge asymmetry generates a voltage. A 1 cm cube of quartz with 500 lb (2 kN) of correctly applied pressure upon it, can produce 12,500 V of electricity.

Piezoelectricity is the ability of certain crystals to produce a voltage when subjected to mechanical stress. The word is derived from the Greek *piezein*, which means to squeeze or press. The effect is reversible; piezoelectric crystals, subject to an externally applied voltage, can change shape by a small amount. The effect is of the order of nanometres, but nevertheless finds useful applications such as the production and detection of sound, generation of high voltages, electronic frequency generation, and ultrafine focusing of optical assemblies.

In a piezoelectric crystal, the positive and negative electrical charges are separated, but symmetrically distributed, so that the crystal overall is electrically neutral. When a stress is applied, this symmetry is disturbed, and the charge asymmetry generates a voltage. A 1 cm cube of quartz with 500 lb (2 kN) of correctly applied pressure upon it, can produce 12,500 V of electricity.

Sensors

- To detect sound, e.g. piezoelectric microphones (sound waves bend the piezoelectric material, creating a changing voltage) and piezoelectric pickups for electrically amplified guitars.
- Piezoelectric elements are also used in the generation of sonar waves. Piezoelectric microbalances are used as very sensitive chemical and biological sensors.
- Piezoelectric elements are used in electronic drum pads to detect the impact of the drummer's sticks.

Mainfold Absolute Pressure (MAP) Sensor

The manifold absolute pressure sensor is a variable resistor used to monitor the difference in pressure between the intake manifold at outside atmosphere.

This information is used by the engine computer to monitor engine load (vacuum drops when the engine is under load or at wide open throttle). When the engine is under load, the computer may alter spark timing and the fuel mixture to improve performance and emissions.

Knock sensor

Engine knock is characteristic of an uncontrolled combustion process and can cause engine damage. That is prevented by a knock sensor, a noise sensor near the engine.

The Knock Sensor is a Piezo Electric device that when you stress it, a voltage is produced. It senses knock and transmits information to the electronic engine management control unit. This influences process control in the engine, for example timing and fuel injection until knock is eliminated.

Knock is frequently caused by fuel that does not comply with the required minimum quality.

Electrical / Electronic & Mechanical Sensors

Objectives : At the end of this lesson you shall be able to

- types of electrical / electronic sensors
- types of mechanical sensors
- explain types of sensors.

Sensors

A sensor is a device that measures a particular characteristic of an object or system. Some sensors are purely mechanical, but most sensors are electronic, returning a voltage signal that can be converted into a useful engineering unit. Sensors take advantage of the mechanical or electrical response of its components to relate the response to a relevant quantity. Engineers use sensors in test and monitoring applications, but homeowners interact with sensors every day. Automobiles are filled with sensors, from the engine to the airbag.

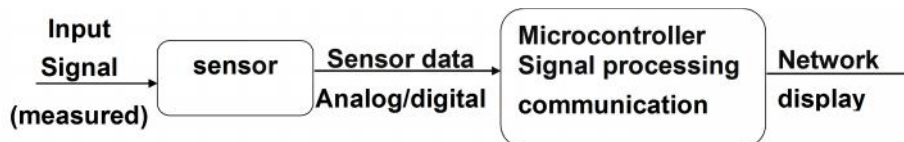
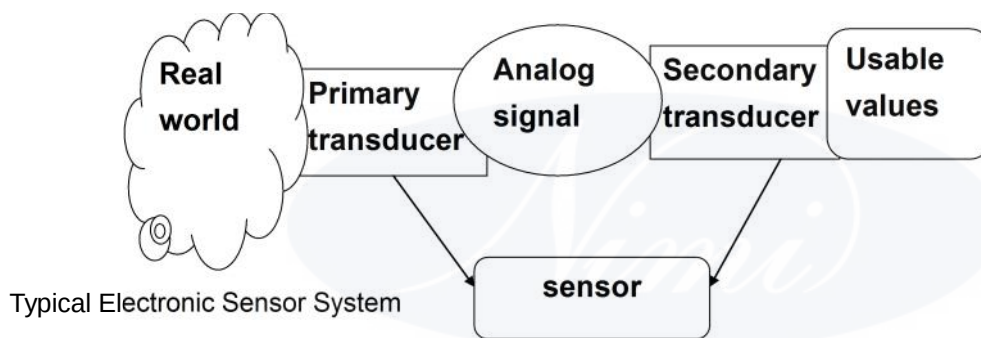
Electrical / Electronic Sensors

convert desired parameter into electrically measurable signal

General Electronic Sensor

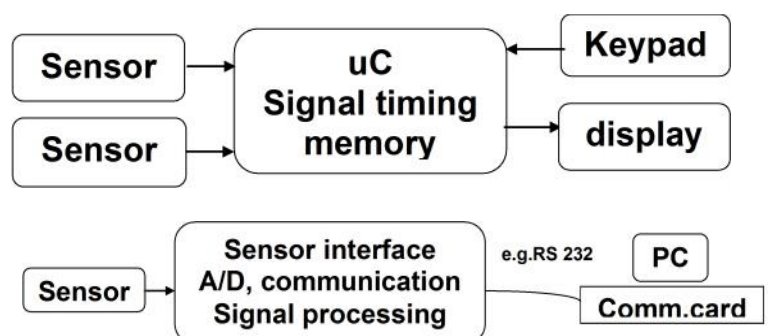
primary transducer: changes "real world" parameter into electrical signal

secondary transducer : converts electrical signal into analog or digital values

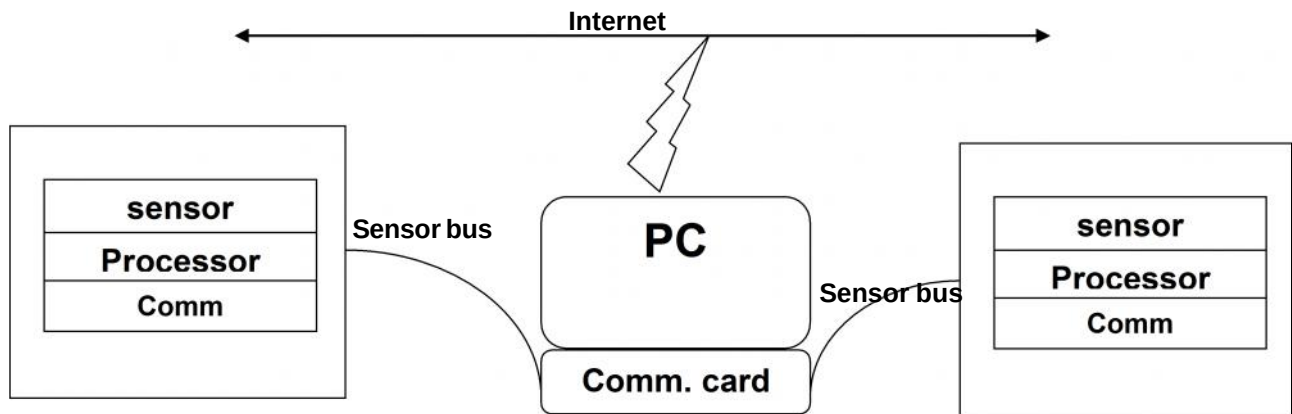


Example Electronic Sensor Systems

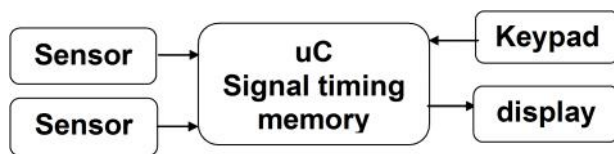
- Components vary with application
 - digital sensor within an instrument
- microcontroller
 - signal timing
 - data storage
 - analog sensor analyzed by a PC



Multiple sensors displayed over internet



Connecting sensors to microcontrollers



- Analog
- many microcontrollers have a built - in A/D
 - 8-bit to 12-bit common
 - many have multi-channel A/D inputs
- Digital
- Serial I/O
 - use serial I/O part, store in memory to analyze
 - synchronous (with clock)
- Must match byte format, stop/start bits, parity check, etc
 - asynchronous (no clock) : more common for comm. than data
- Must match baud rate and bit width, transmission protocol, etc
- **Frequency encoded**
 - use timing port, measure pulse width or pulse frequency
- Interface through lab view or user-generated code
 - Communication links common for sensors
- asynchronous serial comm
 - Universal asynchronous receive and transmit (UART)
 - 1 receive line + 1 transmit line. nodes must match baud rate & protocol
 - RS232 serial port on PCs uses UART format (but at +/-12v)
 - Can buy a chip to convert from UART to RS232
- **Synchronous serial comm**
 - Serial peripheral interface (SPI)
 - 1 clock + 1 bidirectional data + 1 chip select / enable
 - I²C = Inter integrated circuit bus
 - Designed by Philips for comm. inside TVS, used in several commercial sensor systems
- IEC 61451 : Sensor comm. standard
 - Several different sensor comm. protocols for different applications

Connecting smart sensors to PC/Network

- “smart sensor”= sensor with built-in signal processing & communication
- e.g., combining a “dumb sensor” and a microcontroller
 - Data acquisition cards (DAQ)
- PC card with analog and digital I/O

Working with Relays

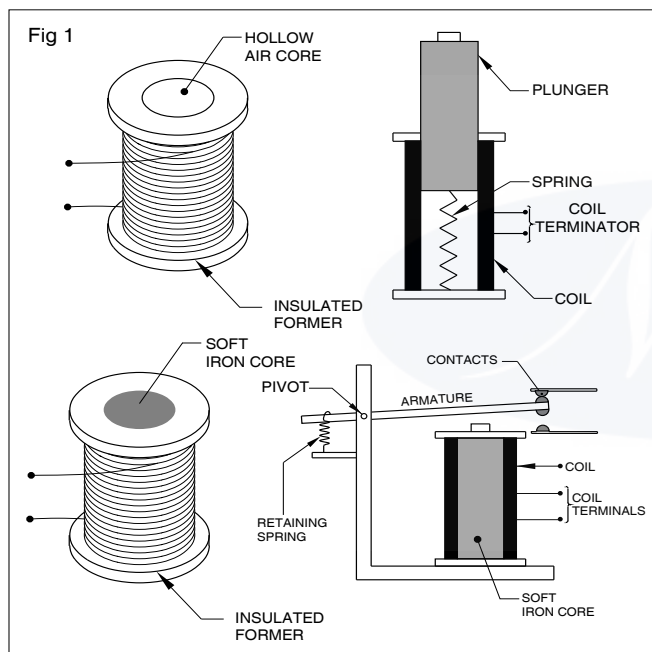
Objectives: At the end of this lesson you shall be able to

- state the difference between solenoid and relays
- explain the general function of relays
- parts of a typical relay
- describe terms Pull-in and Drop-out
- explain the functions of relay parts
- explain the terms Operating current, Holding current, Release current
- list the classification of relays
- list the specifications of a relay coil
- list the different types of contacts in relays
- list a few common relay contact and spring defects.

INTRODUCTION

In addition to solenoids, one other most popular application of electromagnets is in what are called electromagnetic relays.

Important similarities and differences between a solenoid and a relay is illustrated in Fig 1.

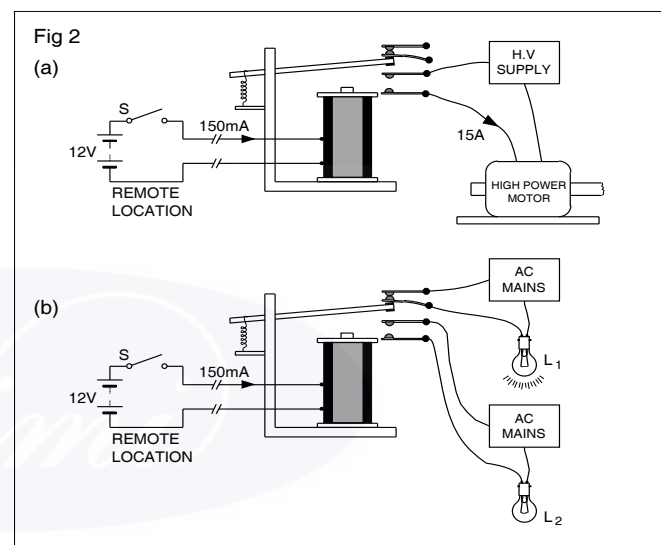


Electromagnetic relays

The term relay was used for the first time, to describe an invention made by Samuel Morse in 1836. The device invented by Morse was a Telegraph Amplifying Electromagnetic Device. 'This device enabled a small current flowing in a coil to switch-ON a large current in another circuit, and thus helped in relaying of telegraph signals.

In any application, the object of a relay is generally to act as a remote switch or as a electrical multiplier switch. This means, a relay enables a comparatively weak current to bring into operation a much stronger current or currents as shown in Fig 2a and Fig 2b.

In Fig 2a, a 150mA relay current can switch-ON or switch-OFF a high power motor which draws 15 Amperes. In Fig 2b, by operating the relay, multiple circuits can be switched-ON or switched-OFF.



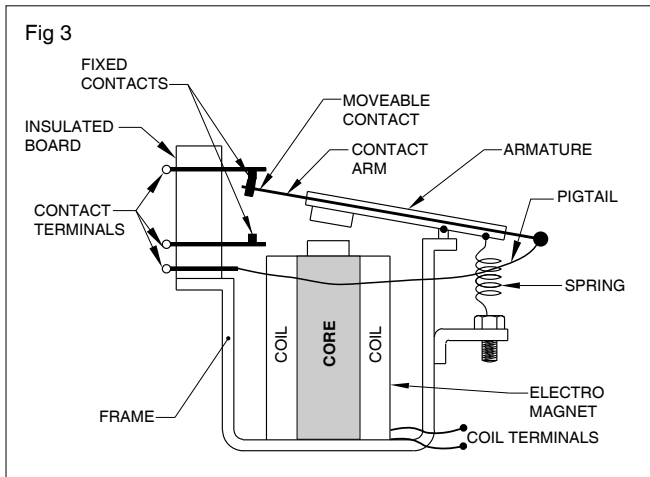
Construction and operation of a simple relay

Electromagnetic relay is basically a switch or a combination of switches operated by magnetic force generated by a current flowing through a coil.

Essentially, a typical relay shown in Fig 3 consists of the following parts;

- an electromagnet comprising of a core and coil
- a movable armature, pivoted and held in tension by a spring
- a set of contacts
- a frame to mount all these components.

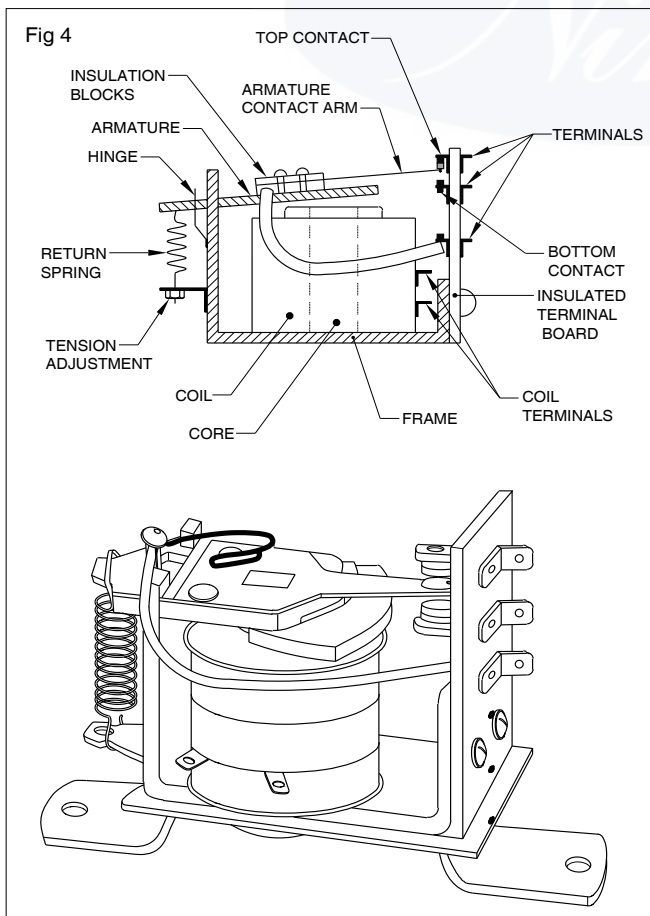
As shown in Fig 3, a typical relay consists of a core surrounded by a coil of wire. This is mounted on a metal frame. The movable part of the relay is the armature. One end of the armature is hinged and connected to a spring. On the armature is mounted a contact arm carrying movable contacts. The fixed relay contacts and its terminals are mounted on an insulated terminal board.



When the relay is OFF or not energized, the contact arm touches the top contact. When the relay is energized by applying voltage to the coil terminals, the metallic armature is attracted. The armature and contact arm assembly move downward so that the contact arm mounted on the armature touches the bottom contact. Thus, the relay is doing the function of a single pole, double throw (SPDT) switch.

On removing the voltage applied to the coil, the spring attached to one end of the armature returns the armature to its original position and the contact arm touches the top contact.

The constructional details of a commercially available, most common type of relay known as a clapper type relay is shown in Fig 4. The working principle of this relay is same as explained for the relay as shown in Fig 3.



Operating delay of relays

When an energizing voltage is applied to a relay coil, the relay does not work instantaneously. It takes some time, usually a few milliseconds to operate. Reasons for this delay are given below:

- Due to inductance of the relay coil, current grows slowly and takes some time to reach the required current value.
- Due to inertia, the armature takes sometime to move from one position to another.

When rated voltage is applied to terminals of a relay coil, the gradual build up of current in the coil is due to the initial opposition to the current flow by the self-inductance of the coil. After some delay, when sufficient magnetization is built up and when the force of attraction is sufficient to overcome the opposition of the tension due to return spring plus, tension of contact springs, the armature is attracted and it closes the relay contacts. The relay is then said to be energized or pulled-in or picked.

Once the relay is energised then, only a small amount of energy is required to maintain it in energized condition. The rest of the electrical energy is wasted as heat.

When the current through the coils falls below a certain value, the relay gets de-energised and the return spring pulls the armature back. This is called as relay drop-out.

From above it can be seen that, very little amount of electrical power is consumed for the switching of relay whereas most of the power is consumed while holding.

PARTS OF A RELAY

Each part of a relay is as important as the other in the overall performance of the relay. Details of the parts of a relay and their purpose are given below;

Frame and core : One of the main function of the relay frame is to provide a base for mounting other relay parts. But, the most important function is, the frame forms a part of the complete magnetic path between the armature and core. The core, frame and armature are made of an easily magnetizable material such as iron.

Hinges : The hinges connect the armature to the frame. A good hinge must be as free from friction as possible. They must also be strong enough to support the weight of the armature and contacts. The hinges must provide low reluctance to the magnetic flux in its path from the core through the frame and the armature.

Return springs : The springs are usually very thin and cannot concentrate any large amount of flux. Spring steel, which has a lower reluctance than other materials acts to retain its magnetism and remain attracted to the core after the relay is de-energised. Springs also have a disadvantage of being stiff and are likely to break after a few operations.

Relay coil : The coil is usually wound on a former and slipped over the magnetic core in the relay frame. This permits easy replacement of damaged coils by new ones.

COIL SPECIFICATIONS

Generally relays are made to operate at different voltages such as, 6, 12, 18, 24, 48, 100 or 240 volts AC or DC. A coil resistance chart is usually given with relays which helps in calculating the coil current and power dissipation. Maximum wattage, maximum permissible temperature and the wattage for satisfactory operation, are specified along with relays.

- Operate current – is the minimum current required to energize a relay.
- Hold current – is the minimum coil current required to continue to hold the relay energized.
- Release current – is the maximum current which releases the relay.

Relay coils are always insulated from the frame of the relay. The electrical resistance between the coil and the body is a measure of the isolation of energising voltage from the ground. Similarly, the electrical resistance between the coil and the contacts is a measure of the electrical isolation between the energising driving and the driven circuits. These resistances will be of the order of hundreds or thousands of megohms.

Relay contacts

The contacts on a relay are the parts that actually perform the electrical switching of the controlled circuits. Also, these contacts are the ones that cause most trouble and require frequent maintenance as compared to any other part of a relay.

Contact materials and design

The relay contacts are made of material which are very good conductors as well as corrosion-resistant.

An arc is created when the contacts open and close. This arc burns and oxidises the contacts. An oxide coating make the contacts either poor conductors or non-conductors. For this reason, contacts are made of silver, palladium and palladium-iridium alloys, gold alloys, gold plated silver, tungsten and alloys of other highly corrosion-resistant materials that do not oxidize easily.

Even with these materials, some oxidation still takes place. To get rid of the oxide, the contacts are designed to have a wiping action. As the contacts close and open, the surfaces rub together. This action rubs off any oxide or dirt which might cause poor contact.

Contacts come in many shapes and sizes, and in a variety of contact arms. These contact arms are generally called contact springs because they maintain good contact pressure.

Size of the contacts determines the current handling capability. The larger the contacts, the more current they can switch without excessive deterioration.

The contact arms or springs are made thick and wide enough to carry the current for which the contacts are rated. They are also made spongy enough to ensure good contact. If the springs are too soft they may vibrate when the relay opens, causing contact bounce when the contacts open and close repeatedly. This bounce can also occur on closing. The bouncing of contacts is always undesirable. Contact debouncing circuits are used to overcome the undesirable effects of contact bouncing in sensitive circuits such as digital electronic circuits.

Switching circuits with switching speed of 10 to 25 msec require gold contacts. Also the contacts should be made of similar metals or alloys to avoid temperature differential which will introduce uncompensated thermocouple voltages.

Types of relay contacts

There are several types of relay contacts of which a few basic types of contacts are shown in Chart 1 at the end of this lesson.

[H.I: Instructor to use transparency of types of contacts and discuss their working.]

The simplest and most common forms of contacts are:

Form A: **Normally Open Contact** (abbreviated as N/O) which closes when the relay gets energized.

Form B: **Normally Closed Contact** (abbreviated as N/C) and the contact gets broken when the relay is energized.

Form C: **Change Over Contact** (Break before make)

Form D: **Make before break contact** - the long arm is contacted first, then it is pushed away from the short arm. In practice this type of contact is very useful as it decides the switching sequence.

Form E: It is basically a type D contact combined with a type B contact.

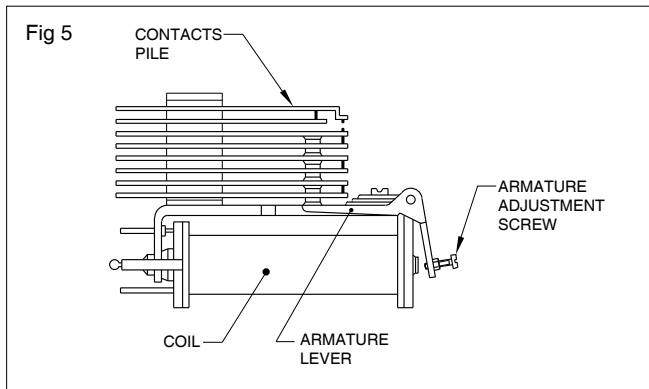
Form F, G, H and J are pile-ups where multiple contacts are made or broken when the relay picks up.

A single relay can have a pile of contacts and operate all of them simultaneously as shown in Fig 5. The contact assembly can contain a variety of contact types shown in Chart 1.

Common relay contact defects

Due to continuous switching the relay contacts and springs develop defects. Some of these defects are illustrated in Chart 2 at the end of this lesson. In most relays, the contact spring assembly can be replaced if they are found defective.

Ask your instructor to show defective contacts and springs and new springs used for replacement.



CLASSIFICATION OF RELAYS

- 1 Relays can be classified according to the type of mechanism used to operate the contacts. Examples clapper, ratchet, stepper, solenoid, meter, reed and thermal relays.
- 2 Relays can also be classified based on their applications as given below;

Contactors - relays with contacts that normally carry 30 or more amperes of current.

Motor starter relays - relays used in motor circuits to control the starting current.

Circuit breakers - relays used as protective devices to interrupt circuits when they become overloaded.

General purpose relays - relays used for general applications.

Telephone type relays - relays used by telephone departments for switching lines and other metering purposes.

Miniature relays - relays used for mounting on printed circuit boards(PCBs) as component of electronic circuit.

3 Relays can also be classified as;

Voltage operated relays - these relays usually will have high coil resistance and draw small amount of current at the rated voltage for its operation. These are usually connected in parallel with the voltage sources.

Current operated relays - these relays usually have low coil resistance and are connected in series with other circuits.

Marginal relays - these relays are used when it is required to operate a relay from a wide range of power such as output of an amplifier which may have a varying input signal strength. These relays permit adjustment of operation and release points within very close limits.

CHART 1 : TYPES OF RELAY CONTACTS

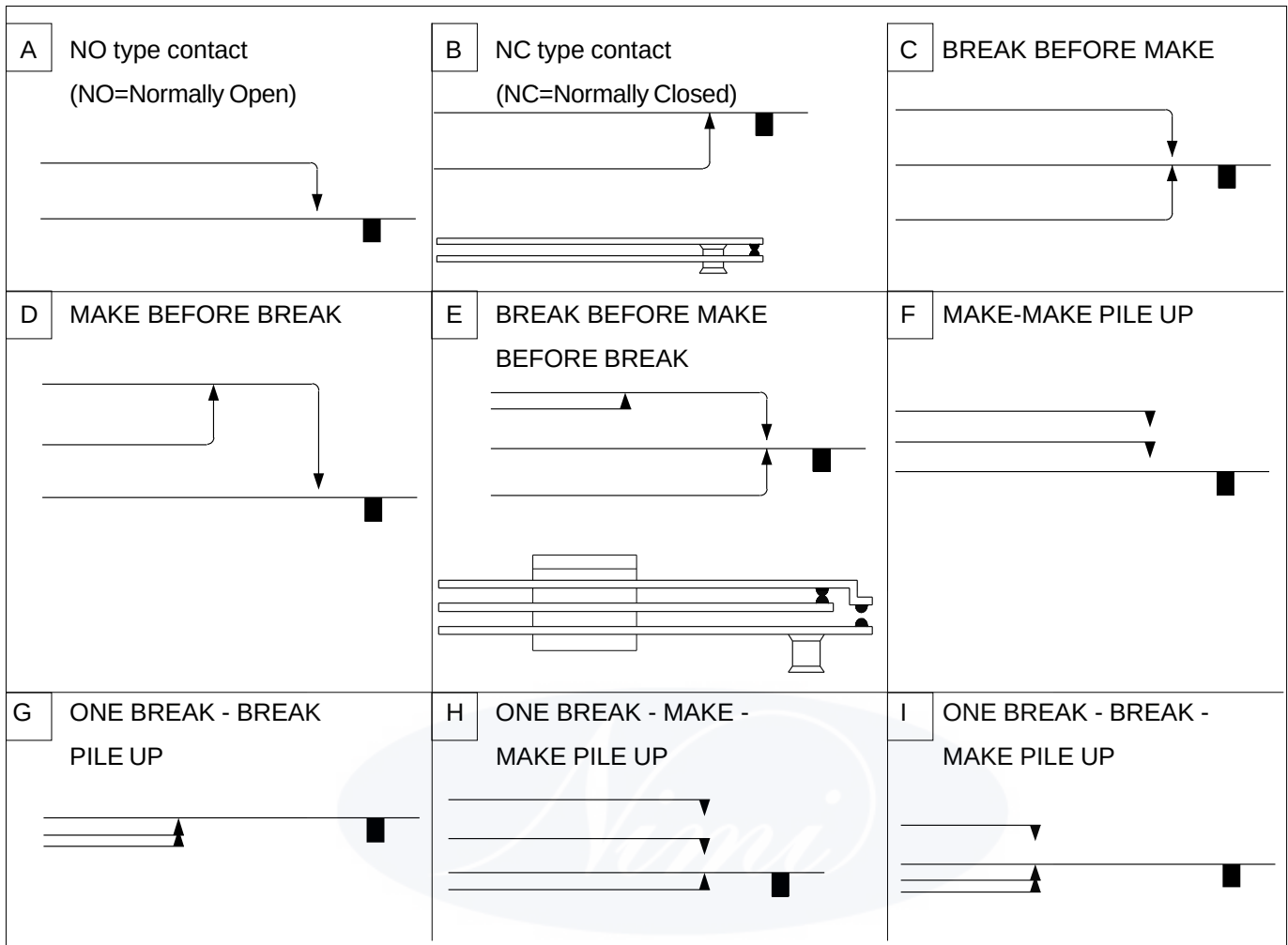
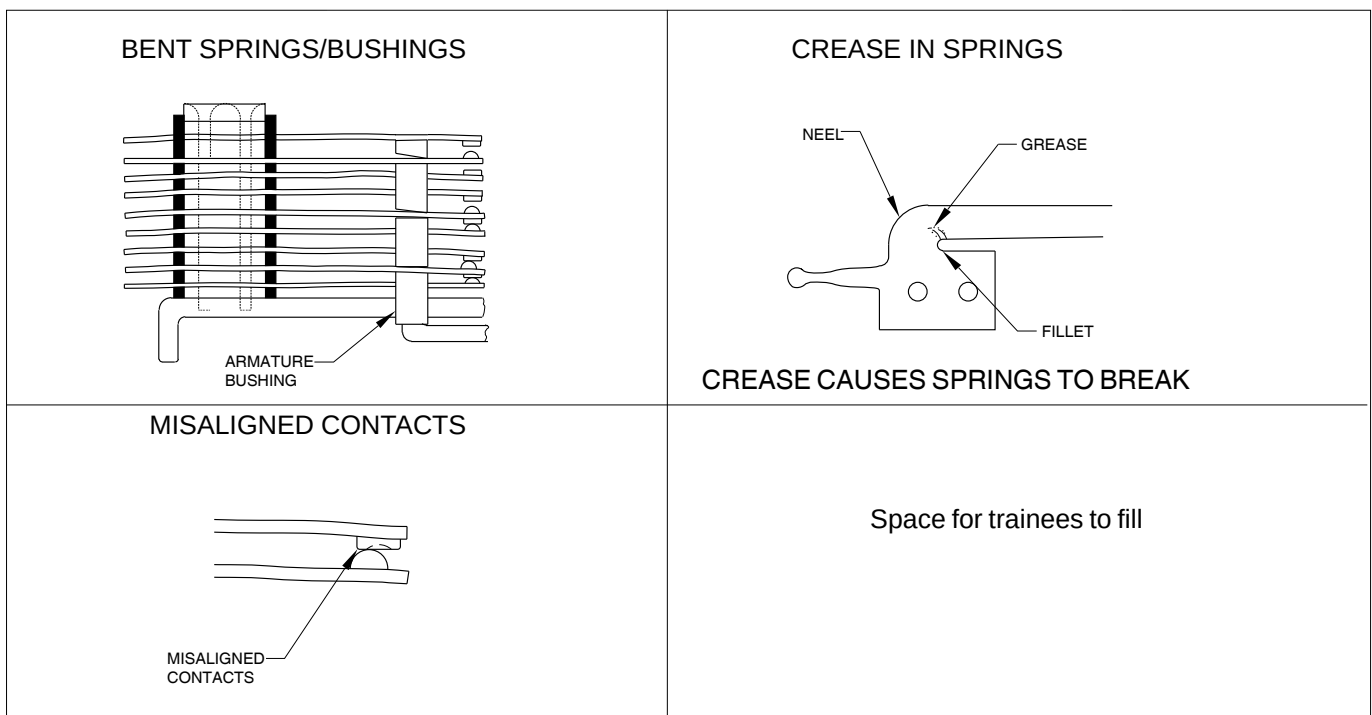


CHART 2 : COMMON RELAY CONTACT/SPRING DEFECTS



Advanced intel microprocessor chips

Objectives: At the end of this lesson you shall be able to

- identify different advanced Intel microprocessor chips.

Processors

The processor also called as microprocessor or CPU is the brain of the PC. It performs all computing tasks and coordinates tasks done by memory, video, disk storage and other system components. CPU is the most complex chip that resides on the motherboard. Almost all modern CPUs come with an execution unit, branch predictor, floating point unit, primary cache etc.

Most processors have an internal cache which stores frequently used data and instructions. Cache is broken up into two classifications, L1 which is the internal cache and L2 which is the external cache.

Each processor have buses that go along with them. The external bus (system bus) allows the processor to connect with other devices, such as expansion cards and slots.

The data bus is used to send and receive data.

The address bus is used to describe memory address locations. Each CPU handles a different width of each of the different types of buses.

Processors speed

The clock that resides in the CPU coordinates all the activities by periodically generating a time reference signal called a clock cycle. It is specified in MHz or GHz. Some instructions require one and others require multiple clock cycles and some CPUs execute multiple instructions in one cycle. This timing depends on the architecture of the CPU, its instruction set and the specific instruction itself. Different processors have their own strengths and weaknesses. Even within a family of processors, the performance differ from one to the other. The first IBM PC CPU, the 8088 ran at 4.77 MHz, today's PC can exceed 1 GHz.

Processor architecture

Over the time, the clock speed of the processors increased. If designers have concentrated only on clock speeds the performance factor would have been only to the extent of speed factor. Recent processors work at 150 times the clock speed of PC/XTs 8088, but provide 1500 times the performance which means major architectural improvements are made in CPUs to get faster. The major architectural features of today's processor are wider data buses and registers, floating point units, pipelining and superscalar architecture.

A list of processors of different generations from a few manufacturers are given in charts 2 & 3.

LGA115x-based Processors Microprocessor 8088

It comes in 40 pin DIP socket from INTEL

Operating Frequency : 5 MHz or 8MHz

This microprocessor was used for many applications which do not require the use of complex calculations. Typical applications include word processing, data processing, games, spreadsheets and some programs which do not require the Windows environment.

Microprocessor 80286

It comes in 68 pin package.

Operating Frequency 12 MHz, 16 MHz, 20 MHz or 25 MHz.

This microprocessor was used for many applications which do not require the use of high efficiency calculations. Typical application includes word processing, data processing, games, spreadsheets and some programs which do not require the Windows environment. Feature with managing functions involving multi tasking and multiuser functions.

Microprocessor 80386SX

It comes in 100 pin mounted from INTEL

Operating Frequency : 25 MHz, 33MHz or 40 MHz

This is a low power version of microprocessor 80386DX. This inexpensive microprocessor ideal for word processing, simple games and Network Node (where the data processing is to quite simple nature) and you can run Windows 3.1

Microprocessor 80386DX:

It comes in 132 pin PGA socket from INTEL

Operating Frequency : 25 MHz, 33MHz or 40 MHz

Useful for most Windows based applications and 2-D imaging. Although, working Windows 3.11 applications is quite convenient with this microprocessor. But if the numeric coprocessor not installed then this inexpensive microprocessor is ideal for Network Server which does not require use of high efficiency numeric calculations. Such applications are rather simple in nature and require the microprocessor basically for managing purposes.

Microprocessor 80486SX

Type of connector : Socket 3 ZIF (Zero Insertion Force) from INTEL

Operating Frequency : 25MHz, 33MHz, 40MHz

Useful for most Windows based applications and 2-d Imaging. But if you are using Windows 95, you will feel that your application is running rather slow. Although working with Windows 3.1 applications is quite convenient with this microprocessor. If the numeric coprocessor is not installed then this inexpensive microprocessor is ideal for Network Server which does not require the use of high efficiency numeric calculations. Such applications are rather simple in nature and require the use of microprocessor basically for managing purposes.

Microprocessor 80486DX/DX-2/DX-4

Type of connector : socket 3 ZIF (Zero Insertion force) from INTEL

Operating Frequency : 25MHz, 33MHz, 40MHz, 50MHz, 66MHz, 80MHz, 100MHz or 120MHz

Comments: A version of 80486 use ful for most Windows based applications and 2D imaging. But if you are using Windows 95 you will feel that your application is running rather slow. Although working with Windows 3.1 applications is quite convenient with this microprocessor.

Microprocessor Pentium

Type of connector : Socket 7 ZIF (Zero Insertion Force) from INTEL

Operating Frequency : 75MHz to 233MHz

Comments : Useful for most Windows based applications. Versions with lower clock speeds might make you feel uncomfortable when you work on 3D imaging. But if you are using Windows 98, you will fee that your application is running rather slow. Although working with Windows 95 applications is quite convenient with this microprocessor. Better still, select a Pentium processor supporting MMX technology.

Microprocessor Pentium Pro

This Microprocessor is manufactured by INTEL

Type of connector : Socket 8 ZIF (Zero Insertion Force) from INTEL

Operating Frequency : 233 MHz to 300 MHz

Comments : Useful for most Windows based applications, Web surfing, and 2 D imaging. Versions with lower clock speeds might make you feel uncomfortable when you work on 3 D imaging. Working with Windows 98 applications is quite convenient with this microprocessor.

Microprocessor Celeron

This microprocessor is manufactured by INTEL

Type of connector : Slot 1 or Socket 370 ZIF

Operating Frequency : 266MHz to 450MHz

Useful for most windows based applications, multimedia and web surfing, 2D imaging. Versions with lower clock speeds might make you feel uncomfortable when you work on 3D imaging.

Microprocessor Pentium II

This microprocessor is manufactured by INTEL

Type of connector : Slot 1

Operating frequency : 300 MHz to 550 MHz

Useful for Multimedia 3D virtual worlds, Video conferencing, Internet, PC imaging and some other new Windows based applications.

Microprocessor MXI (cayenne)

This microprocessor is manufactured by Cyrix

Type of connector : Socket 7

Operating Frequency : 400 MHz upwards

A completely new design. Suitable for Multimedia, 3D virtual worlds, video conferencing, internet, PC imaging and some other new Windows based applications.

Microprocessor K6-3(sharptooth)

This microprocessor is manufactured by AMD (Advance Micro Devices)

Type of connector : Socket 7

Operating Frequency : 350 - 450 MHz

This is an advanced version of the K6-2 microprocessor of the same company. Suitable for Multimedia, 3D virtual Worlds, Video Conferencing, Internet, PC imaging and some other new Windows based applications. This microprocessor is expected to be faster than the Pentium II at the same clock speeds.

Microprocessor Pentium III

This microprocessor is manufactured by INTEL

Type of connector : Slot 1

Operating Frequency : 450 MHz upwards

This microprocessor is an improved version of pentium II and it is very suitable for Multimedia, 3D Virtual worlds, Video conferencing, internet, PC imaging and some other new windows based application.

Microprocessor K7

This microprocessor is manufactured by AMD

Type of connector : Slot A(Alpha EV-6)

Operating Frequency : 500MHz upwards

This is an advanced version of the K6 microprocessor from

the same company. Suitable for Multimedia, 3D virtual worlds, Video conferencing, Internet, PC imaging and some other new windows based applications.

Microprocessor MIII (JALAOENO)

This microprocessor is manufactured by cyrix.

Type of connector : Super 7

Operating Frequency : 600 MHz upwards

A completely new CPU design. Suitable for Multimedia, 3D virtual worlds, Video conferencing, Internet, PC imaging and some other new windows based applications.

Microprocessor IA-64 (MERCD):

This microprocessor is manufactured by INTEL

Type of connector : Slot M

Operating Frequency : 800 MHz upwards

This microprocessor has a totally new internal architecture. This microprocessor is useful for multimedia, 3D virtual worlds, Video conferencing, Internet, PC imaging and some other new windows based applications.

Nowadays processors of speeds above 3 GHz are available from manufacturers like Intel AMD etc.

The following tables summarize the processor chips and their characteristics:

CPU	Data bus (bits)	Address Bus (bits)	Speed (MHz)	Transistors
8088	8	20	4.77	29,000
80286	16	24	8-12	134,000
80386SX	32	24	16-20	275,000
80386DX	32	32	16-33	275,000
80486SX	32	32	16-33	1.185 Million
80486DX	32	32	25-50	1.2 Million
486DX2	32	32	33-66	2 Million
486DX4	32	32	75-100	2.5 Million
Pentium	32	32	60-166	3.3 Million
Pentium Pro	64	32	150-200	5.5 Million
Pentium Pro II	64	64	233-400	7.5 Million

CPUs are mounted in different sockets depending on the size of the chip. Also, 486 chips and newer usually contain heat sinks, with fans, which fit on top of the processor chip. The CPU socket information is summarized in this table:

Socket	Voltage	Pins	Processors	Upgrade
0	3.3	168	486DX	486DX2/486DX4
1	3.3	169	486DX, 486SX	486DX2, 486DX4
2	3.3	238	486DX, 486SX, 486DX2	486DX2/4, Pentium
3	3.3 or 5	237	486DX, 486SX, 486DX2, 486DX4	486DX2/486DX4
4	5	273	60/66 MHz Pentium	Pentium
5	3.3	320	Other Pentium	Pentium
6	3.3	235	486DX4	Pentium
7	3.3	321	Other Pentium	Pentium
8	3.3	387	Pentium Pro	Pentium Pro

There are two types of upgrade methods, ZIF (Zero Insertion Force) and LIF (Low Insertion Force) sockets. ZIF sockets have a mounting bar attached.

Advanced Intel microprocessor chips

Objectives: At the end of this lesson you shall be able to

- identify different advanced Intel microprocessor chips.

Advanced Intel microprocessor chips

Core i3

Core i3 is the new low end of the performance processor line from Intel, following the retirement of the Core 2 brand

The first Core i3 processors were launched on January 7, 2010

The Core i3-3xxM processors are similar to the Core i5-4xx series but running at lower clock speeds and without Turbo Boost. they do not support Error Correction Code (ECC) memory. According to motherboard manufacturer Supermicro, if a Core i3 processor is used with a server chipset platform such as Intel 3400/3420/3450, the CPU supports ECC with UDIMM.

Core i5

The Core i5-5xx mobile processors are named Arrandale and based on the 32 nm Westmere shrink of the Nehalem microarchitecture. Arrandale processors have integrated graphics capability but only two processor cores. They were released in January 2010, together with Core i7-6xx and Core i3-3xx processors based on the same chip. The L3 cache in Core i5-5xx processors is reduced to 3 MB, while the Core i5-6xx uses the full cache and the Core i3-3xx does not support for Turbo Boost. Clarkdale, the desktop version of Arrandale, It has Hyper-Threading enabled and the full 4 MB L3 cache.

Core i5 desktop processors and desktop boards typically do not support ECC memory.

Core i7

Intel Core i7 as an Intel brand name applies to several families of desktop and laptop 64-bit x86-64 processors using the Nehalem, Westmere, Sandy Bridge, Ivy Bridge and Haswell micro architectures. The Core i7 brand targets the business and high-end consumer markets for both desktop and laptop computers, and is distinguished from the Core i3 (entry-level consumer), Core i5 (mainstream consumer), and Xeon (server and workstation) brands.

In each of the first three microarchitecture generations of the brand, Core i7 has family members using two distinct system-level architectures, and therefore two distinct sockets (for example, LGA 1156 and LGA 1366 with Nehalem). In each generation, the highest-performing Core i7 processors use the same socket and QPI-based architecture as the low-end Xeon processors of that generation, while lower-performing Core i7 processors use the same socket and PCIe/DMI/FDI architecture as the Core i5.

"Core i7" is a successor to the Intel Core 2 brand.

Processor Markings

All Intel® Desktop Processors have markings inscribed on the top of the processor package. The most recent Intel® Desktop Processor packages and markings are described here.

Some of these markings include the serial number of the processor which is required for warranty processing. The serial number of a processor is typically a combination of the finished process order (FPO) number and the assembly test process order (ATPO) number. Always provide both numbers if they are available for your processor. If the ATPO number is not present, provide the FPO number.

LGA115x-based Processors



Line 1 Copyright Marks

Line 2 Trademark and Branding + Processor Family (Example: Intel® Core™ i5 Processor)

Line 3 sSpec # + CPU Speed

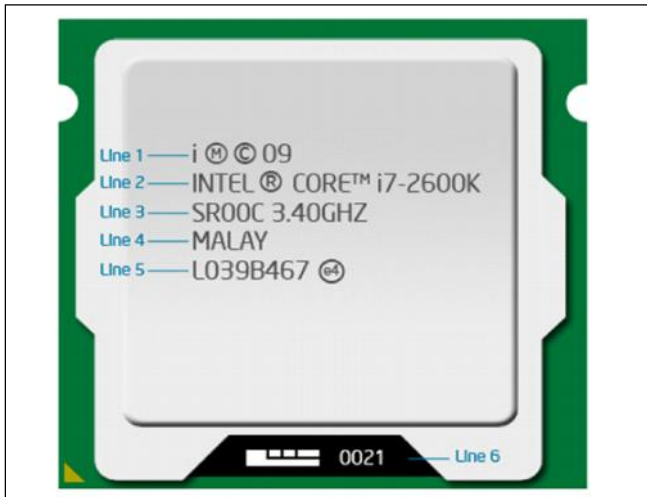
Line 4 Country of Manufacture

Line 5 Finished Process Order (FPO)

Line 6 Partial Assembly Test Process Order (ATPO)
Serial Number = Line 5 + Line 6 (Example: L070Q228-1234)

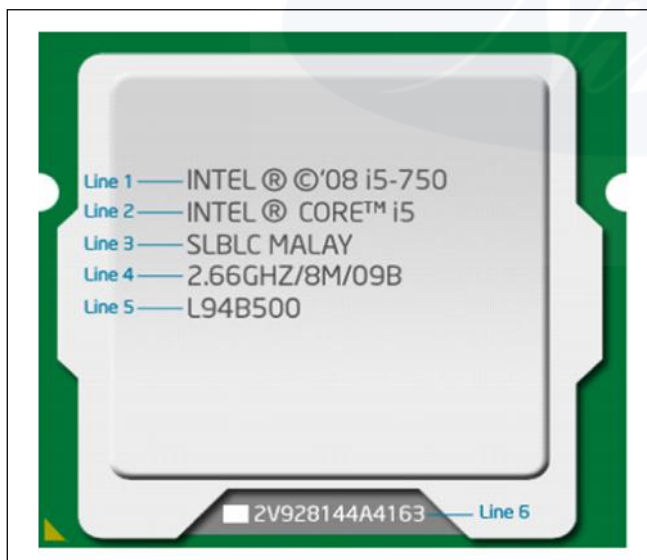
sSpec = Line 3 (Example: SR28F)

LGA 1155 - Based Processors



Line 1 Copyright Marks + Platform Compatibility Guide
 Line 2 Trademark and Branding + Processor Family (Example: Intel® Core™ i7 Processor) + Processor Number (Example: 2600K)
 Line 3 sSpec # + CPU Speed
 Line 4 Country of Manufacture
 Line 5 Finished Process Order (FPO)
 Line 6 Partial Serial Number
 Serial Number = Line 5 + Line 6 (Example: L039B467-0021)
 sSpec = Line 3 (Example: SR00C)

LGA 1156 - Based Processors



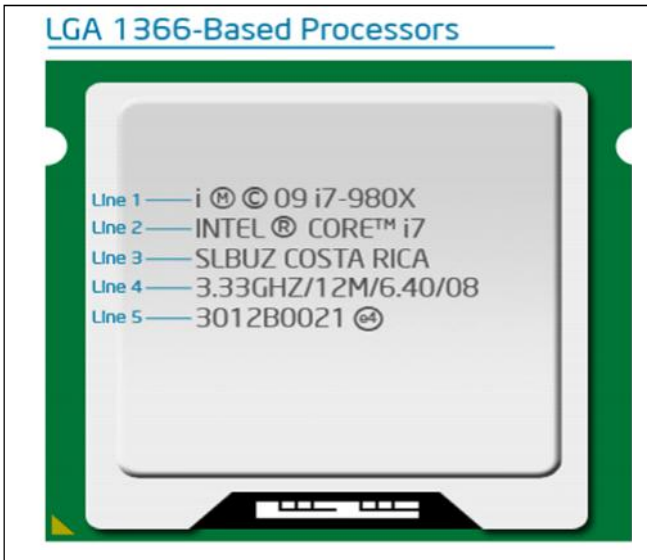
Line 1 Processor number (Example: Intel® Core™ i5-750 Processor)
 Line 2 Processor Family (Example: Intel® Core™ i5 Desktop Processor)
 Line 3 sSpec # and Manufacturing Country
 Line 4 CPU speed/L2 cache size/Platform Compatibility Guide
 Line 5 Finished Process Order (FPO)
 Line 6 Assembly Test Process Order (ATPO) + Serial Number
 Serial Number = Line 5 + Line 6 (Example: L94B500-2V928144A4163)
 sSpec = Line 3 (Example: SLBLC)

LGA2011-based Processors + LGA1366-based Processors



Line 1 Copyright Marks
 Line 2 Trademark and Branding + Processor Family + Processor Number (Example: Intel® Core™ i7 Processor-3930K)
 Line 3 sSpec # + CPU Speed
 Line 4 Country of Manufacture
 Line 5 Finished Process Order (FPO)
 Line 6 Assembly Test Process Order (ATPO) + Serial Number
 Serial Number = Line 5 (Example: 3134B507)
 sSpec = Line 3 (Example: SR0H9)

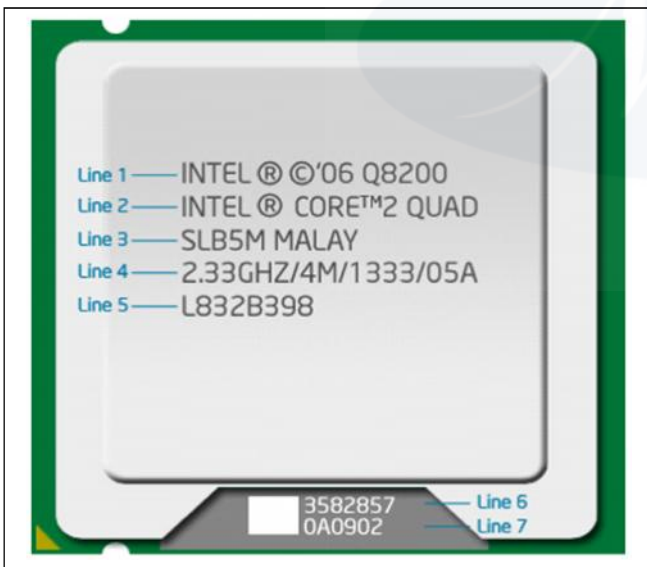
LGA 1366-Based Processors



- Line 1 Processor number (Example: Intel® Core™2 Quad Processor Q8200)
 - Line 2 Processor Family (Example: Intel® Core™2 Quad Processor)
 - Line 3 sSpec # and Manufacturing Country
 - Line 4 CPU speed/L2 cache size/Bus speed/ Platform Compatibility Guide
 - Line 5 Finished Process Order (FPO)
 - Line 6 Assembly Test Process Order (ATPO)
 - Line 7 Partial Serial Number
- Serial Number = Line 5 + Line 6 + Line 7 (Example: L832B398-35828570A0902)
- sSpec = Line 3 (Example: SLB5M)

- Line 1 Copyright Marks + Processor Number (Example: 980X)
 - Line 2 Trademark and Branding + Processor Family (Example: Intel® Core™ i7 Processor)
 - Line 3 sSpec # + Country of Manufacture
 - Line 4 CPU Speed/Level2 Cache/Intel QPI Speed/ Platform Compatibility Guide
 - Line 5 Finished Process Order (FPO)
- Serial Number = Line 5 (Example: 3012B0021)
- sSpec = Line 3 (Example: SLBUZ)

LGA775-based Processors



Advanced microprocessor chips other than from Intel

Objectives : At the end of this lesson you shall be able to

- **identify different advanced non-Intel microprocessor chips.**
-

One of Intel's major competitors is Advanced Micro Devices (AMD). AMD has developed a family of processors that compete with Pentium-class processors. In this section, I provide an overview of some of the characteristics of the AMD processors.

K6

The AMD K6 processor was designed to compete with the original Intel Pentium. The K6 has 64K of L1 cache, supports MMX technology, and has built-in branch prediction techniques. This processor has 321 pins, which means that it will fit into a Socket 7-supported motherboard.

K6-2

The K6-2 processor was designed to compete with the Pentium II chip. It has 64K of L1 cache and 256K of L2 cache. The K6-2 also supports dynamic execution, MMX technology, and superscalar design.

The K6-2 has added 3DNow! technology, comprising a number of additional instructions integrated into the chip to improve 3-D graphics applications. The K6-2 chip also uses a 100 MHz motherboard speed, which is a big improvement over the 60/66 MHz motherboard speed that the original Pentiums used.

The K6-2 has 321 pins, which means that it will fit into a Socket 7-supported motherboard.

K6-III

The K6-III processor is designed to compete with the Pentium III chip. This chip shares many of the features of the K6-2, including a 100 MHz system bus. One of its features was the tri-level cache. Not only can it take advantage of an L1 and L2 cache but also an L3 cache that can be included on the motherboard.

Athlon

The AMD Athlon chip has 128K of L1 cache and 512K of L2 cache. It supports improved dynamic execution, MMX technology, and 3DNow! technology. The Athlon chip runs at speeds of up to 1.2 GHz and is designed to run on a 200 MHz system bus speed.

Unlike the K6-2 and K6-III, the Athlon is not a PGA-packaged chip that supports Socket 7. It uses its own socket type - Slot A, so called because the processor is packaged as an SEC. The Slot A socket is not compatible with Intel's Slot 1, which means that users have to purchase a motherboard designed for the Athlon chip.

remember Later versions of the Athlon moved to the PGA package that has 462 pins. These PGA chips are placed in Socket A.

Athlon XP

After the Athlon chip was produced, Intel created the Pentium 4 chip. So AMD wanted to create a competing chip for the Pentium 4: namely, the Athlon XP. The Athlon XP is packaged as a PGA with 462 pins and is placed in Socket A. The Athlon XP runs at 2 GHz or more and contains 128K of L1 cache and 512K of L2 cache.

technical stuff AMD markets these processors a little differently. Instead of labeling the processor with its speed, AMD labels it with its competitor's speed. For example, the Athlon XP 1800+ is rated at 1.6 GHz but runs as fast as the Intel 1.8 GHz processor.

Duron

AMD wanted to create a processor that competed with each version of the Intel processors. So, if the Athlon XP competes with the Pentium 4, what competes with the Celeron? You guessed it - the Duron.

The Duron has 128K of L1 cache and 64K of L2 cache. This processor is packaged as a PGA with 462 pins, which means that it, too, goes into Socket A.

Opteron

Just like the Duron was built to compete with the Intel Celeron, AMD created the Opteron to compete with Intel 64-bit Itanium processors. The Opteron runs at about 1.8 GHz and contains 128K of L1 cache and 1MB of L2 cache.

The Opteron is packaged with a micro-PGA, which is made up of 940 pins and is placed in Socket 940. One of the major differences between the Opteron and the Itanium is that the Itanium cannot run 32-bit applications. AMD decided that the Opteron would run in a 32-bit or 64-bit mode, thus allowing it to run 32-bit applications.

Athlon 64 and Athlon 64 X2

The Athlon 64 - the successor to the Athlon XP - is the AMD 64-bit processor for desktop systems. The Athlon 64, with 128K of L1 cache and at least 512K of L2 cache built into the processor, is designed to compete with the Pentium 4.

Although the Athlon 64 is a 64-bit processor, it has been designed to be backward compatible; it can run 32-bit code. The Athlon 64 X2 is the dual-core version of the Athlon 64. The Athlon 64 family processors fit into a number of sockets: Socket 754, Socket 940, and the AM2 socket.

For The ExamKnow that the Sempron processor is a low-end version of the Athlon 64 and has replaced the Duron as the AMD low-end processor. The Sempron was originally a 32-bit processor with reduced cache size, but current versions are 64-bit with the cache size reduced.

Phenom and Phenom II

The Phenom and Phenom II pick up after the Athlon 64 and are the AMD triple- and quad-core processors. The Phenom, designed for desktop systems, uses the code names Toliman for the triple-core version and Agena for the quad-core version.

The Phenom processor, which comes with 128K of L1 cache and 512KB of L2 cache per core, also has 2MB of shared L3 cache. The processor fits in the AMD AM2+ socket and runs between approximately 1.8 GHz and 2.6 GHz.

The Phenom II also sits in the AM2+ socket and increases the shared L3 cache to 6MB! There were issues with the Phenom running on Windows Vista, but those bugs have been fixed for the Phenom

II. The Phenom II has a triple-core version, code-named Heka, that runs between 2.6 GHz and 2.8 GHz. The quad-core version that is named Deneb runs between 2.5 GHz and 3.0 GHz.

Turion 64 and Turion 64 X2

The Turion 64 processor is the AMD 64-bit mobile processor for use in laptop computers. The Turion processor used to be called the Athlon Mobile 64, but AMD has moved to the new label of Turion 64 for its mobile processors. The Turion 64 is a single-core processor, and the Turion 64 X2 is the dual-core processor.

Sensors

Sensors are devices that are frequently used to detect and respond to electrical or optical signals. A Sensor converts the physical parameter (for example: temperature, blood pressure, humidity, speed, etc.) into a signal which can be measured electrically. For example the temperature sensor. The mercury in the glass thermometer expands and contracts the liquid to convert the measured temperature which can be read by a viewer on the calibrated glass tube.

Criteria to choose a Sensor

There are certain features which have to be considered when we choose a sensor. They are as given below:

- 1 **Accuracy**
- 2 **Environmental condition** : usually has limits for temperature/ humidity

- 3 **Range** : Measurement limit of sensor
- 4 **Calibration** : Essential for most of the measuring devices as the readings changes with time
- 5 **Resolution** : Smallest increment detected by the sensor
- 6 **Cost**
- 7 **Repeatability** : The reading that varies is repeatedly measured under the same environment

Classification of Sensors

The sensors are classified into the following criteria:

- 1 Primary Input quantity
- 2 Transduction principles (Using physical and chemical effects)
- 3 Material and Technology
- 4 Property
- 5 Application

Transduction principle is the fundamental criteria which are followed for an efficient approach. Usually, material and technology criteria are chosen by the development engineering group.

Classification based on property is as given below:

- **Temperature** : Thermistors, thermocouples, RTD's, IC and many more.
- **Pressure** : Fibre optic, vacuum, elastic liquid based manometers, LVDT, electronic.
- **Flow** : Electromagnetic, differential pressure, positional displacement, thermal mass, etc.
- **Level Sensors** : Differential pressure, ultrasonic radio frequency, radar, thermal displacement, etc.
- **Proximity and displacement** : LVDT, photoelectric, capacitive, magnetic, ultrasonic.
- **Biosensors** : Resonant mirror, electrochemical, surface Plasmon resonance, Light addressable potentio-metric.
- **Image** : Charge coupled devices, CMOS
- **Gas and chemical** : Semiconductor, Infrared, Conductance, Electrochemical.
- **Acceleration** : Gyroscopes, Accelerometers.
- **Others** : Moisture, humidity sensor, Speed sensor, mass, Tilt sensor, force, viscosity.

The Bio-sensors group are the new optical technology based sensors. CMOS Image sensors have low resolution as compared to charge coupled devices. CMOS has the advantages of small size, cheap, less power consumption and hence are better substitutes for Charge coupled devices. Accelerometers are independently grouped because of their vital role in future applications like aircraft, automobiles, etc and in fields of videogames, toys, etc.

Classification based on Application is as given below:

- Industrial process control, measurement and automation
- **Non-industrial use** : Aircraft, Medical products, Automobiles, Consumer electronics, other type of sensors.

Classified based on power or energy supply requirement of the sensors:

- **Active Sensor** : Sensors that require power supply are called as Active Sensors. Example: LiDAR (Light detection and ranging), photoconductive cell.
- **Passive Sensor** : Sensors that do not require power supply are called as Passive Sensors. Example: Radiometers, film photography.
- In the current and future applications, sensors can be classified into groups as follows:

- **Accelerometers** : These are based on the Micro Electro Mechanical sensor technology. They are used for patient monitoring which includes pace makers and vehicle dynamic systems.
- **Biosensors** : These are based on the electrochemical technology. They are used for food testing, medical care device, water testing, and biological warfare agent detection.
- **Image Sensors** : These are based on the CMOS technology. They are used in consumer electronics, biometrics, traffic and security surveillance and PC imaging.
- **Motion Detectors** : These are based on the Infra Red, Ultrasonic, and Microwave / radar technology. They are used in videogames and simulations, light activation and security detection.



