“Introduction to electrical engineering”.

In Section 1 of this course you will cover these topics:

- Introduction
- Resistive Circuits
- Inductance And Capacitance
- Transients

Topic : Introduction

Topic Objective:

At the end of this topic student would be able to:

- Recognize interrelationships between electrical engineering and other fields of science and engineering.
- List the major subfields of electrical engineering.
- List several important reasons for studying electrical engineering.
- Define current, voltage, and power, including their units.
- Calculate power and energy and determine whether energy is supplied or absorbed by a circuit element.
- State and apply Kirchhoffs current and voltage laws.
- Recognize series and parallel connections.
- Identify and describe the characteristics of voltage and current sources.
- State and apply Ohms law.
- Solve for currents, voltages, and powers in simple circuits.

Definition/Overview:

Electrical Engineering: Electrical engineering, sometimes referred to as electrical and electronic engineering, is a field of engineering that deals with the study and application of electricity, electronics and electromagnetism.

Circuits: An electronic circuit is a closed path formed by the interconnection of electronic components through which an electric current can flow.

Currents: Electric current is the flow (movement) of electric charge. The SI unit of electric current is the ampere. Electric current is measured using an ammeter.

Voltages: Electrical tension (or voltage after its SI unit, the volt) is the difference of electrical potential between two points of an electrical or electronic circuit, expressed in volts.

Power: Electric power is defined as the rate at which electrical energy is transferred by an electric circuit. The SI unit of power is the watt.

Energy: In physics and other sciences, energy is a scalar physical quantity that is a property of objects and systems which is conserved by nature. Energy is often defined as the ability to do work.

Key Points:

1. Electronic Engineering

1.1. Overview

Electrical engineering, sometimes referred to as electrical and electronic engineering, is a field of engineering that deals with the study and application of electricity, electronics and electromagnetism. The field first became an identifiable occupation in the late nineteenth century after commercialization of the electric telegraph and electrical power supply. It now covers a range of subtopics including power, electronics, control systems, signal processing and telecommunications.

Electrical engineering may or may not encompass electronic engineering. Where a distinction is made, usually outside of the United States, electrical engineering is considered to deal with the problems associated with large-scale electrical systems such as power transmission and motor control, whereas electronic engineering deals with the study of small-scale electronic systems including computers and integrated circuits. Alternatively, electrical engineers are usually concerned with using electricity to transmit energy, while electronic engineers are concerned with using electricity to transmit information.

1.2. History

Electricity has been a subject of scientific interest since at least the early 17th century. The first electrical engineer was probably William Gilbert who designed the versorium: a device that detected the presence of statically charged objects. He was also the first to draw a clear distinction between magnetism and static electricity and is credited with establishing the term electricity. In 1775 Alessandro Volta's scientific experimentations devised the electrophorus, a device that produced a static electric charge, and by 1800 Volta developed the voltaic pile, a forerunner of the electric battery.

However, it was not until the 19th century that research into the subject started to intensify. Notable developments in this century include the work of Georg Ohm, who in 1827 quantified the relationship between the electric current and potential difference in a conductor, Michael Faraday, the discoverer of electromagnetic induction in 1831, and James Clerk Maxwell, who in 1873 published a unified theory of electricity and magnetism in his treatise Electricity and Magnetism.

During these years, the study of electricity was largely considered to be a subfield of physics. It was not until the late 19th century that universities started to offer degrees in electrical engineering. The Darmstadt University of Technology founded the first chair and the first faculty of electrical engineering worldwide in 1882. In 1883 Darmstadt University of Technology and Cornell University introduced the world's first courses of study in electrical engineering, and in 1885 the University College London founded the first chair of electrical engineering in the United Kingdom. The University of Missouri subsequently established the first department of electrical engineering in the United States in 1886.

During this period, the work concerning electrical engineering increased dramatically. In 1882, Edison switched on the world's first large-scale electrical supply network that provided 110 volts direct current to fifty-nine customers in lower Manhattan. In 1887, Nikola Tesla filed a number of patents related to a competing form of power distribution known as alternating current. In the following years a bitter rivalry between Tesla and Edison, known as the "War of Currents", took place over the preferred method of distribution. AC eventually replaced DC for generation and power distribution, enormously extending the range and improving the safety and efficiency of power distribution.

The efforts of the two did much to further electrical engineeringTesla's work on induction motors and polyphase systems influenced the field for years to come, while Edison's work on telegraphy and his development of the stock ticker proved lucrative for his company, which ultimately became General Electric. However, by the end of the 19th century, other key figures in the progress of electrical engineering were beginning to emerge.

1.3. Modern developments

During the development of radio, many scientists and inventors contributed to radio technology and electronics. In his classic UHF experiments of 1888, Heinrich Hertz transmitted (via a spark-gap transmitter) and detected radio waves using electrical equipment. In 1895, Nikola Tesla was able to detect signals from the transmissions of his New York lab at West Point (a distance of 80.4 km / 49.95 miles). In 1897, Karl Ferdinand Braun introduced the cathode ray tube as part of an oscilloscope, a crucial enabling technology for electronic television. John Fleming invented the first radio tube, the diode, in 1904. Two years later, Robert von Lieben and Lee De Forest independently developed the amplifier tube, called the triode. In 1895, Guglielmo Marconi furthered the art of hertzian wireless methods. Early on, he sent wireless signals over a distance of one and a half miles. In December 1901, he sent wireless waves that were not affected by the curvature of the Earth. Marconi later transmitted the wireless signals across the Atlantic between Poldhu, Cornwall, and St. John's, Newfoundland, a distance of 2,100 miles (3,400 km). In 1920 Albert Hull developed the magnetron which would eventually lead to the development of the microwave oven in 1946 by Percy Spencer. In 1934 the British military began to make strides towards radar (which also uses the magnetron) under the direction of Dr Wimperis, culminating in the operation of the first radar station at Bawdsey in August 1936.

In 1941 Konrad Zuse presented the Z3, the world's first fully functional and programmable computer. In 1946 the ENIAC (Electronic Numerical Integrator and Computer) of John Presper Eckert and John Mauchly followed, beginning the computing era. The arithmetic performance of these machines allowed engineers to develop completely new technologies and achieve new objectives, including the Apollo missions and the NASA moon landing.

The invention of the transistor in 1947 by William B. Shockley, John Bardeen and Walter Brattain opened the door for more compact devices and led to the development of the integrated circuit in 1958 by Jack Kilby and independently in 1959 by Robert Noyce. In 1968 Marcian Hoff invented the first microprocessor at Intel and thus ignited the development of the personal computer. The first realization of the microprocessor was the Intel 4004, a 4-bit processor developed in 1971, but only in 1973 did the Intel 8080, an 8-bit processor, make the building of the first personal computer, the Altair 8800, possible.

1.4. Education

Electrical engineers typically possess an academic degree with a major in electrical engineering. The length of study for such a degree is usually four or five years and the completed degree may be designated as a Bachelor of Engineering, Bachelor of Science, Bachelor of Technology or Bachelor of Applied Science depending upon the university. The degree generally includes units covering physics, mathematics, computer science, project management and specific topics in electrical engineering. Initially such topics cover most, if not all, of the sub-disciplines of electrical engineering. Students then choose to specialize in one or more sub-disciplines towards the end of the degree.

Some electrical engineers also choose to pursue a postgraduate degree such as a Master of Engineering/Master of Science (MEng/MSc), a Master of Engineering Management, a Doctor of Philosophy (PhD) in Engineering, an Engineering Doctorate (EngD), or an Engineer's degree. The Master and Engineer's degree may consist of either research, coursework or a mixture of the two. The Doctor of Philosophy and Engineering Doctorate degrees consist of a significant research component and are often viewed as the entry point to academia. In the United Kingdom and various other European countries, the Master of Engineering is often considered an undergraduate degree of slightly longer duration than the Bachelor of Engineering.

1.5. Practicing engineers

In most countries, a Bachelor's degree in engineering represents the first step towards professional certification and the degree program itself is certified by a professional body. After completing a certified degree program the engineer must satisfy a range of requirements (including work experience requirements) before being certified. Once certified the engineer is designated the title of Professional Engineer (in the United States, Canada and South Africa), Chartered Engineer (in India, the United Kingdom, Ireland and Zimbabwe), Chartered Professional Engineer (in Australia and New Zealand) or European Engineer (in much of the European Union).

The advantages of certification vary depending upon location. For example, in the United States and Canada "only a licensed engineer may seal engineering work for public and private clients". This requirement is enforced by state and provincial legislation such as Quebec's Engineers Act. In other countries, such as Australia, no such legislation exists. Practically all certifying bodies maintain a code of ethics that they expect all members to abide by or risk expulsion. In this way these organizations play an important role in maintaining ethical standards for the profession. Even in jurisdictions where certification has little or no legal bearing on work, engineers are subject to contract law. In cases where an engineer's work fails he or she may be subject to the tort of negligence and, in extreme cases, the charge of criminal negligence. An engineer's work must also comply with numerous other rules and regulations such as building codes and legislation pertaining to environmental law.

Professional bodies of note for electrical engineers include the Institute of Electrical and Electronics Engineers (IEEE) and the Institution of Engineering and Technology (IET) (which was formed by the merging of the Institution of Electrical Engineers (IEE) and the Institution of Incorporated Engineers (IIE). The IEEE claims to produce 30% of the world's literature in electrical engineering, has over 360,000 members worldwide and holds over 3,000 conferences annually. The IET publishes 21 journals, has a worldwide membership of over 150,000, and claims to be the largest professional engineering society in Europe. Obsolescence of technical skills is a serious concern for electrical engineers. Membership and participation in technical societies, regular reviews of periodicals in the field and a habit of continued learning are therefore essential to maintaining proficiency.

In countries such as Australia, Canada and the United States electrical engineers make up around 0.25% of the labor force. Outside of these countries, it is difficult to gauge the demographics of the profession due to less meticulous reporting on labour statistics. However, in terms of electrical engineering graduates per-capita, electrical engineering graduates would probably be most numerous in countries such as Taiwan, Japan, India and South Korea.

1.6. Sub-Disciplines

Electrical engineering has many sub-disciplines, the most popular of which are listed below. Although there are electrical engineers who focus exclusively on one of these sub-disciplines, many deal with a combination of them. Sometimes certain fields, such as electronic engineering and computer engineering, are considered separate disciplines in their own right.

1.6.1. Power

Power engineering deals with the generation, transmission and distribution of electricity as well as the design of a range of related devices. These include transformers, electric generators, electric motors, high voltage engineering and power electronics. In many regions of the world, governments maintain an electrical network called a power grid that connects a variety of generators together with users of their energy. Users purchase electrical energy from the grid, avoiding the costly exercise of having to generate their own. Power engineers may work on the design and maintenance of the power grid as well as the power systems that connect to it. Such systems are called on-grid power systems and may supply the grid with additional power, draw power from the grid or do both. Power engineers may also work on systems that do not connect to the grid, called off-grid power systems, which in some cases are preferable to on-grid systems. The future includes Satellite controlled power systems, with feedback in real time to prevent power surges and prevent blackouts.

1.6.2. Control

Control engineering focuses on the modeling of a diverse range of dynamic systems and the design of controllers that will cause these systems to behave in the desired manner. To implement such controllers electrical engineers may use electrical circuits, digital signal processors, microcontrollers and PLCs (Programmable Logic Controllers). Control engineering has a wide range of applications from the flight and propulsion systems of commercial airliners to the cruise control present in many modern automobiles. It also plays an important role in industrial automation.

Control engineers often utilize feedback when designing control systems. For example, in an automobile with cruise control the vehicle's speed is continuously monitored and fed back to the system which adjusts the motor's power output accordingly. Where there is regular feedback, control theory can be used to determine how the system responds to such feedback.

1.6.3. Electronics

Electronic engineering involves the design and testing of electronic circuits that use the properties of components such as resistors, capacitors, inductors, diodes and transistors to achieve a particular functionality. The tuned circuit, which allows the user of a radio to filter out all but a single station, is just one example of such a circuit.

Prior to the second world war, the subject was commonly known as radio engineering and basically was restricted to aspects of communications and radar, commercial radio and early television. Later, in post war years, as consumer devices began to be developed, the field grew to include modern television, audio systems, computers and microprocessors. In the mid to late 1950s, the term radio engineering gradually gave way to the name electronic engineering.

Before the invention of the integrated circuit in 1959, electronic circuits were constructed from discrete components that could be manipulated by humans. These discrete circuits consumed much space and power and were limited in speed, although they are still common in some applications. By contrast, integrated circuits packed a large numberoften millionsof tiny electrical components, mainly transistors, into a small chip around the size of a coin.

1.6.4. Microelectronics

Microelectronics engineering deals with the design and microfabrication of very small electronic circuit components for use in an integrated circuit or sometimes for use on their own as a general electronic component. The most common microelectronic components are semiconductor transistors, although all main electronic components (resistors, capacitors, inductors) can be created at a microscopic level.

Microelectronic components are created by chemically fabricating wafers of semiconductors such as silicon (at higher frequencies, compound semiconductors like gallium arsenide and indium phosphide) to obtain the desired transport of electronic charge and control of current. The field of microelectronics involves a significant amount of chemistry and material science and requires the electronic engineer working in the field to have a very good working knowledge of the effects of quantum mechanics.

1.6.5. Signal Processing

Signal processing deals with the analysis and manipulations of signals. Signals can be either analog, in which case the signal varies continuously according to the information, or digital, in which case the signal varies according to a series of discrete values representing the information. For analog signals, signal processing may involve the amplification and filtering of audio signals for audio equipment or the modulation and demodulation of signals for telecommunications. For digital signals, signal processing may involve the compression, error detection and error correction of digitally sampled signals.

1.6.6. Telecommunications

Telecommunications engineering focuses on the transmission of information across a channel such as a coax cable, optical fiber or free space. Transmissions across free space require information to be encoded in a carrier wave in order to shift the information to a carrier frequency suitable for transmission, this is known as modulation. Popular analog modulation techniques include amplitude modulation and frequency modulation. The choice of modulation affects the cost and performance of a system and these two factors must be balanced carefully by the engineer.

Once the transmission characteristics of a system are determined, telecommunication engineers design the transmitters and receivers needed for such systems. These two are sometimes combined to form a two-way communication device known as a transceiver. A key consideration in the design of transmitters is their power consumption as this is closely related to their signal strength. If the signal strength of a transmitter is insufficient the signal's information will be corrupted by noise.

1.6.7. Instrumentation Engineering

Instrumentation engineering deals with the design of devices to measure physical quantities such as pressure, flow and temperature. The design of such instrumentation requires a good understanding of physics that often extends beyond electromagnetic theory. For example, radar guns use the Doppler effect to measure the speed of oncoming vehicles. Similarly, thermocouples use the Peltier-Seebeck effect to measure the temperature difference between two points.

Often instrumentation is not used by itself, but instead as the sensors of larger electrical systems. For example, a thermocouple might be used to help ensure a furnace's temperature remains constant. For this reason, instrumentation engineering is often viewed as the counterpart of control engineering.

1.6.8. Computers

Computer engineering deals with the design of computers and computer systems. This may involve the design of new hardware, the design of PDAs or the use of computers to control an industrial plant. Computer engineers may also work on a system's software. However, the design of complex software systems is often the domain of software engineering, which is usually considered a separate discipline. Desktop computers represent a tiny fraction of the devices a computer engineer might work on, as computer-like architectures are now found in a range of devices including video game consoles and DVD players.

1.6.9. Related Disciplines

Mechatronics is an engineering discipline which deals with the convergence of electrical and mechanical systems. Such combined systems are known as electromechanical systems and have widespread adoption. Examples include automated manufacturing systems, heating, ventilation and air-conditioning systems and various subsystems of aircraft and automobiles.

The term mechatronics is typically used to refer to macroscopic systems but futurists have predicted the emergence of very small electromechanical devices. Already such small devices, known as micro electromechanical systems (MEMS), are used in automobiles to tell airbags when to deploy, in digital projectors to create sharper images and in inkjet printers to create nozzles for high definition printing. In the future it is hoped the devices will help build tiny implantable medical devices and improve optical communication.

Biomedical engineering is another related discipline, concerned with the design of medical equipment. This includes fixed equipment such as ventilators, MRI scanners and electrocardiograph monitors as well as mobile equipment such as cochlear implants, artificial pacemakers and artificial hearts.

2. Circuits

An electronic circuit is a closed path formed by the interconnection of electronic components through which an electric current can flow. The electronic circuits may be physically constructed using any number of methods. Breadboards, perfboards or stripboards are common for testing new designs. Mass-produced circuits are typically built using a printed circuit board (PCB) that is used to mechanically support and electrically connect electronic components.

Electronic circuits can display highly complex behaviors, even though they are governed by the same laws of physics as simpler circuits. It can usually be categorized as analog, discrete, or mixed-signal (a combination of analog and discrete) electronic circuits.

2.1. Analog Circuits

Analog electronic circuits are those in which signals may vary continuously with time to correspond to the information being represented. Electronic equipment like voltage amplifiers, power amplifiers, tuning circuits, radios, and televisions are largely analog (with the exception of their control sections, which may be digital, especially in modern units).

The basic units of analog circuits are passive (resistors, capacitors, inductors, and recently memristors) and active (independent power sources and dependent power sources). Components such as transistors may be represented by a model containing passive components and dependent sources. Another classification is to take impedance and independent sources and operational amplifier as basic electronic components; this allows us to model frequency dependent negative resistors, gyrators, negative impedance converters, and dependent sources as secondary electronic components. There are two main types of circuits: series and parallel. A string of Christmas lights is a good example of a series circuit: if one goes out, they all do. In a parallel circuit, each bulb is connected to the power source separately, so if one goes out the rest still remain shining.

2.2. Discrete Circuits

In digital electronic circuits, electric signals take on discrete values, which are not dependent upon time, to represent logical and numeric values. These values represent the information that is being processed. The transistor is one of the primary components used in discrete circuits, and combinations of these can be used to create logic gates. These logic gates may then be used in combination to create a desired output from an input.

Larger circuits may contain several complex components, such as FPGAs or Microprocessors. These along with several other components may be interconnected to create a large circuit that operates on large amount of data. Examples of electronic equipment which use digital circuits include digital wristwatches, calculators, PDAs, and microprocessors.

2.3. Mixed-Signal Circuits

Mixed-signal or hybrid circuits contain elements of both analog and digital circuits. Examples include comparators, timers, PLLs, ADCs (analog-to-digital converters), and DACs (digital-to-analog converters).

3. Kirchhoff's Circuit Laws

Kirchhoff's circuit laws are a pair of laws that deal with the conservation of charge and energy in electrical circuits, and were first described in 1845 by Gustav Kirchhoff. Widely used in electrical engineering, they are also called Kirchhoff's rules or simply Kirchhoff's laws. Both circuit rules can be directly derived from Maxwell's equations, but Kirchhoff preceded Maxwell and instead generalized work by Georg Ohm.

3.1. Kirchhoff's Current Law (KCL)

This law is also called Kirchhoff's first law, Kirchhoff's point rule, Kirchhoff's junction rule (or nodal rule), and Kirchhoff's first rule. The principle of conservation of electric charge implies that: At any point in an electrical circuit where charge density is not changing in time, the sum of currents flowing towards that point is equal to the sum of currents flowing away from that point.

An analogy to this principle is: Two rivers that converge and then later break up into separate rivers. The principle states that the sum of the water flowing in the two upstream rivers is equal to the sum of the water flowing in the two downstream rivers.

A charge density changing in time would mean the accumulation of a net positive or negative charge, which typically cannot happen to any significant degree because of the strength of electrostatic forces: the charge buildup would cause repulsive forces to disperse the charges.

Another way to look at the effect of changing charge density on Kirchhoff's first rule is to go back to the rivers analogy again. Whenever there is a change in the charge density imagine someone going downstream and adding or removing a bucket of water from any of the two downstream rivers. This will change the rate of flow downstream.

However, a charge build-up can occur in a capacitor, where the charge is typically spread over wide parallel plates, with a physical break in the circuit that prevents the positive and negative charge accumulations over the two plates from coming together and cancelling. In this case, the sum of the currents flowing into one plate of the capacitor is not zero, but rather is equal to the rate of charge accumulation. However, if the displacement current dD/dt is included, Kirchhoff's current law once again holds. (This is only required if one wants to apply the current law within the capacitor. In circuit analyses, however, the capacitor as a whole is typically treated as a unit, in which case the ordinary current law holds since the net charge is always zero.)

More technically, Kirchhoff's current law can be found by taking the divergence of Ampre's law with Maxwell's correction and combining with Gauss's law, yielding:

This is simply the charge conservation equation (in integral form, it says that the current flowing out of a closed surface is equal to the rate of loss of charge within the enclosed volume). Kirchhoff's current law is equivalent to the statement that the divergence of the current is zero, true for time-invariant ρ, or always true if the displacement current is included with J. A matrix version of Kirchhoff's current law is the basis of most circuit simulation software, such as SPICE.

3.2. Kirchhoff's Voltage Law (KVL)

This law is also called Kirchhoff's second law, Kirchhoff's loop (or mesh) rule, and Kirchhoff's second rule. The directed sum of the electrical potential differences around any closed circuit must be zero. KVL may also be stated as the algebraic sum of various potential drops across an electrical circuit is equal to the electromotive force acting on the circuit". This statement is equivalent to the statement that a single-valued electric potential can be assigned to each point in the circuit (in the same way that any conservative vector field can be represented as the gradient of a scalar potential).

(This could be viewed as a consequence of the principle of conservation of energy. Otherwise, it would be possible to build a perpetual motion machine that passed a current in a circle around the circuit).

Considering that electric potential is defined as a line integral over an electric field, Kirchhoff's voltage law can be expressed equivalently as

This states that the line integral of the electric field around closed loop C is zero.

Or

(Round a loop/mesh)

This is a simplification of Faraday's law of induction for the special case where there is no fluctuating magnetic field linking the closed loop. In the presence of a changing magnetic field the electric field is not conservative and it cannot therefore define a pure scalar potentialthe line integral of the electric field around the circuit is not zero. This is because energy is being transferred from the magnetic field to the current (or vice versa). In order to "fix" Kirchhoff's voltage law for circuits containing inductors, an effective potential drop, or electromotive force (emf), is associated with each inductance of the circuit, exactly equal to the amount by which the line integral of the electric field is not zero by Faraday's law of induction.

3.3. Ohm's Law

Ohm's law applies to electrical circuits; it states that the current through a conductor between two points is directly proportional to the potential difference (i.e. voltage drop or voltage) across the two points, and inversely proportional to the resistance between them.

The mathematical equation that describes this relationship is: I = {V} / {R}

where I is the current in amperes, V is the potential difference in volts,and R is a circuit parameter called the resistance (measured in ohms, also equivalent to volts per ampere). The potential difference is also known as the voltage drop, and is sometimes denoted by U, E or emf (electromotive force) instead of V I is from the German Intensitt meaning "intensity". The law was named after the German physicist Georg Ohm, who, in a treatise published in 1827, described measurements of applied voltage and current through simple electrical circuits containing various lengths of wire. He presented a slightly more complex equation than the one above to explain his experimental results. The above equation is the modern form of Ohm's law.

The resistance of most resistive devices (resistors) is constant over a large range of values of current and voltage. When a resistor is used under these conditions, the resistor is referred to as an ohmic device (or an ohmic resistor) because a single value for the resistance suffices to describe the resistive behavior of the device over the range. When sufficiently high voltages are applied to a resistor, forcing a high current through it, the device is no longer ohmic because its resistance, when measured under such electrically stressed conditions, is different (typically greater) from the value measured under standard conditions.

Ohm's law, in the form above, is an extremely useful equation in the field of electrical/electronic engineering because it describes how voltage, current and resistance are interrelated on a "macroscopic" level, that is, commonly, as circuit elements in an electrical circuit. Physicists who study the electrical properties of matter at the microscopic level use a closely related and more general vector equation, sometimes also referred to as Ohm's law, having variables that are closely related to the I, V and R scalar variables of Ohm's law, but are each functions of position within the conductor.

Topic : Resistive Circuits

Topic Objective:

At the end of this topic student would be able to:

- Solve circuits (i.e., find currents and voltages of interest) by combining resistances in series and parallel.
- Apply the voltage-division and current-division principles.

- Solve circuits by the node-voltage technique.
- Solve circuits by the mesh-current technique.
- Find Thvenin and Norton equivalents and apply source transformations.
- Apply the superposition principle.
- Draw the circuit diagram and state the principles of operation for the Wheatstone bridge.

Definition/Overview:

Resistive Circuit: A resistive circuit is an electrical circuit designed to use resistance as a means of controlling the behavior of the electrical current in the circuit. A light bulb is an example of a useful resistive circuit. Many devices such as floor heaters, electric stoves, ovens, clothes dryers, etc. use resistive circuits to generate heat.

Key Points:

1. Series-Parallel Circuit Analysis

When solving for voltage, current, and resistance in a series-parallel circuit, follow the rules which apply to the series part of the circuit, and follow the rules which apply to the parallel part of the circuit. Solving these circuits can be simplified by reducing the circuit to a single equivalent resistance circuit, and redrawing the circuit in simplified form. The circuit is then called an equivalent circuit.

It consists of follwing steps:

- Begin by locating a combination of resistances that are in series or parallel. Often the place to start is farthest from the source.

- Redraw the circuit with the equivalent resistance for the combination found in step 1.
- Repeat steps 1 and 2 until the circuit is reduced as far as possible. Often (but not always) we end up with a single source and a single resistance.
- Solve for the currents and voltages in the final equivalent circuit.

1.1. Voltage Division

A series connected circuit is often referred to as a voltage divider circuit. The source voltage equals the total of all voltage drops across the series connected resistors. The voltage dropped across each resistor is proportional to the resistance value of that resistor. Larger resistors experience larger drops, while smaller resistors experience smaller drops. The voltage divider formula allows you to calculate the voltage drop across any resistor without having to first solve for the current. The voltage divider formula is:

where VX = voltage dropped across selected resistor

RX = selected resistors value

RT = total series circuit resistance

VS = source or applied voltage

1.2. Current Division

Sometimes it is necessary to find the individual branch currents in a parallel circuit when only resistance and total current are known. When only two branches are involved, the current in one branch will be some fraction of IT. The resistance in each circuit can be used to divide the total current into fractional currents in each branch. This process is known as current division.

Using Ohms law, we can derive a formula (called the current divider) that can be used to calculate the current through any branch of a multiple-branch parallel circuit:

Where,

Ix = current to be calculated, in branch x

RT= total resistance

Rx = resistance in branch x

IT = total current

2. Node-Voltage Analysis

In electrical engineering node-voltage analysis, or the branch current method is a method of determining the voltage (potential difference) between "nodes" (points where elements or branches connect) in an electrical circuit. It is used to solve for the voltages and currents at any point in a circuit without working through many individual KCL or KVL rules.

It simplifies the number of equations that would have been developed through KCL and KVL to just the equations that describe the voltage and current according to a node within the circuit diagram. It uses KCL by realizing that the total current entering a node must equal the total current leaving a node. KVL is used to find the currents by using the voltage drop between the node and other nodes to find the current required to go through these paths to create the voltage drop. In other words, nodal analysis uses KVL to find the voltages that would create currents that would satisfy KCL.

This method is very powerful as many different circuit elements can be modeled. Active circuit elements such as operational amplifiers can be added to the analysis. These elements can be as simple or complicated as desired to achieve the fidelity needed in the simulation. For example: A number of different transistor models are available that can be used in the nodal analysis. The only requirement is that the elements are linear.

2.1. Method

- Label all the nodes in the circuit (e.g. 1, 2, 3...), and select one to be the "reference node." It is usually most convenient to select the node with the most connections as the reference node.
- Assign a variable to represent the voltage of each labeled node (e.g. V1, V2, V3...). The values of these variables, when calculated, will be relative to the reference node (i.e. the reference node will be 0V).
- If there is a voltage source between any node and the reference node, by Kirchhoff's voltage law, the voltage at that node is the same as the voltage source's. For example, if there is a 40 V source between node 1 and the reference node, node 0, V1 = 40 V.
- Note any voltage sources between two nodes. These two nodes form a supernode. By Kirchhoff's voltage law, the voltage difference at these two nodes is the same as the voltage source. For example, if there is a 60 V source between node 1 and node 2, V1 − V2 = 60 V.
- For all remaining nodes, write a Kirchhoff's current law equation for the currents leaving each node, using the terminal equations for circuit elements to relate circuit elements to currents. For example, if there is a 10 Ω resistor between nodes 2 and 3, a 1 A current source between nodes 2 and 4 (leaving node 2), and a 20 Ω resistor between nodes 2 and 5, the KCL equation would be (V2 − V3)/10 + 1 + (V2 − V5)/20 = 0 A.
- For all sets of nodes that form a supernode, write a KCL equation, as in the last step for all currents leaving the supernode, i.e. sum the currents leaving the nodes of the supernode. For example, if there is a 60 V source between nodes 1 and 2, nodes 1 and 2 form a supernode. If there is a 40 Ω resistor between nodes 1 and the reference node, a 2

A current source between nodes 1 and 3 (leaving node 3), and a 30 Ω resistor between nodes 2 and 4, the KCL equation would be (V1 − 0)/40 + (−2) + (V2 − V4)/30 = 0 A.

- The KCL and KVL equations form a system of simultaneous equations that can be solved for the voltage at each node.

3. Mesh Analysis

Mesh analysis (sometimes referred to as loop analysis or mesh current method) is a method that is used to solve planar circuits for the voltage and currents at any place in the circuit. Planar circuits are circuits that can be drawn on a plane with no wires overlapping each other. Mesh analysis uses Kirchhoffs voltage law to solve these planar circuits. The advantage of using mesh analysis is that it creates a systematic approach to solving planar circuits and eliminates the number of equations needed to solve the circuit for all of the voltages and currents.

3.1. Mesh Currents and Essential Meshes

Mesh analysis works by arbitrarily assigning mesh currents in the essential meshes. An essential mesh is a loop in the circuit that does not contain any other loop. When looking at a circuit schematic, the essential meshes look like a window pane. Figure 1 labels the essential meshes with one, two, and three. Once the essential meshes are found, the mesh currents need to be labeled.

A mesh current is a current that loops around the essential mesh. The mesh current might not have a physical meaning but it is used to set up the mesh analysis equations. When assigning the mesh currents it is important to have all the mesh currents loop in the same direction. This will help prevent errors when writing out the equations. The convention is to have all the mesh currents looping in a clockwise direction.

3.2. Choosing the Mesh Currents

When several mesh currents flow through one element, we consider the current in that element to be the algebraic sum of the mesh currents. Sometimes it is said that the mesh currents are defined by soaping the window panes.

3.3. Setting Up the Equations

After labeling the mesh currents, one only needs to write one equation per mesh in order to solve for all the currents in the circuit. These equations are the sum of the voltage drops in a complete loop of the mesh current. For other than current and voltage sources, the voltage drops will be the impedance of the electronic component multiplied by the mesh current in that loop. It is important to note that if a component exists between two essential meshes, the component's voltage drop will be the impedance of the component times the present mesh current minus the neighboring mesh current.

If a voltage source is present within the mesh loop, the voltage at the source is either added or subtracted depending on if it is a voltage drop or a voltage rise in the direction of the mesh current. For a current source that is not contained between two meshes, the mesh current will take the positive or negative value of the current source depending on if the mesh current is in the same or opposite direction of the current source. The following is the same circuit from above with the equations needed to solve for all the currents in the circuit.

Once the equations are found, the system of linear equations can be solved by using any technique to solve linear equations.

3.4. Mesh-Current Analysis

If necessary, redraw the network without crossing conductors or elements. Then define the mesh currents flowing around each of the open areas defined by the network. For consistency, we usually select a clockwise direction for each of the mesh currents, but this is not a requirement.

Write network equations, stopping after the number of equations is equal to the number of mesh currents. First, use KVL to write voltage equations for meshes that do not contain current sources. Next, if any current sources are present, write expressions for their currents in terms of the mesh currents. Finally, if a current source is common to two meshes, write a KVL equation for the supermesh.

If the circuit contains dependent sources, find expressions for the controlling variables in terms of the mesh currents. Substitute into the network equations, and obtain equations having only the mesh currents as unknowns.

Put the equations into standard form. Solve for the mesh currents by use of determinants or other means.

Use the values found for the mesh currents to calculate any other currents or voltages of interest.

4. Thvenin and Norton Equivalent Circuits

4.1. Thvenin's theorem

In electrical circuit theory, Thvenin's theorem for linear electrical networks states that any combination of voltage sources, current sources and resistors with two terminals is electrically equivalent to a single voltage source V and a single series resistor R. For single frequency AC systems the theorem can also be applied to general impedances, not just resistors. This theorem states that a circuit of voltage sources and resistors can be converted into a Thvenin equivalent, which is a simplification technique used in circuit analysis. The Thvenin equivalent can be used as a good model for a power supply or battery (with the resistor representing the internal impedance and the source representing the electromotive force). The circuit consists of an ideal voltage source in series with an ideal resistor.

4.2. Norton's Theorem

Norton's theorem is an extension of Thvenin's theorem. Norton's theorem for electrical networks states that any collection of voltage sources, current sources, and resistors with two terminals is electrically equivalent to an ideal current source, I, in parallel with a single resistor, R. For single-frequency AC systems the theorem can also be applied to general impedances, not just resistors. The Norton equivalent is used to represent any network of linear sources and impedances, at a given frequency. The circuit consists of an ideal current source in parallel with an ideal impedance (or resistor for non-reactive circuits).

4.3. Calculating the Thvenin Equivalent

To calculate the equivalent circuit, one needs a resistance and some voltage - two unknowns. And so, one needs two equations. These two equations are usually obtained by using the following steps, but any conditions one places on the terminals of the circuit should also work:

- Calculate the output voltage, VAB, when in open circuit condition (no load resistor - meaning infinite resistance). This is VTh.
- Calculate the output current, IAB, when those leads are short circuited (load resistance is 0). RTh equals VTh divided by this IAB.

The equivalent circuit is a voltage source with voltage VTh in series with a resistance RTh.

Step 2 could also be thought of like this:

- Now replace voltage sources with short circuits and current sources with open circuits.
- Replace the load circuit with an imaginary ohm meter and measure the total resistance, R, "looking back" into the circuit. This is RTh.

The Thvenin-equivalent voltage is the voltage at the output terminals of the original circuit. When calculating a Thvenin-equivalent voltage, the voltage divider principle is often useful, by declaring one terminal to be Vout and the other terminal to be at the ground point.

The Thvenin-equivalent resistance is the resistance measured across points A and B "looking back" into the circuit. It is important to first replace all voltage- and current-sources with their internal resistances. For an ideal voltage source, this means replace the voltage source with a short circuit. For an ideal current source, this means replace the current source with an open circuit. Resistance can then be calculated across the terminals using the formulae for series and parallel circuits.

5. Maximum Power Transfer

Sometimes in engineering we are asked to design a circuit that will transfer the maximum power to a load from a given source. According to the maximum power transfer theorem, a load will receive maximum power from a source when its resistance (RL) is equal to the internal resistance (RI) of the source. If the source circuit is already in the form of a Thevenin or Norton equivalent circuit (a voltage or current source with an internal resistance), then the solution is simple. If the circuit is not in the form of a Thevenin or Norton equivalent circuit, we must first use Thevenins or Nortons theorem to obtain the equivalent circuit.

6. Superposition Theorem

The superposition theorem states that in a linear circuit with several sources, the current and voltage for any element in the circuit is the sum of the currents and voltages produced by each source acting independently. To calculate the contribution of each source independently, all the other sources must be removed and replaced without affecting the final result. When removing a voltage source, its voltage must be set to zero, which is equivalent to replacing the voltage source with a short circuit. When removing a current source, its current must be set to zero, which is equivalent to replacing the current source with an open circuit.

When you sum the contributions from the sources, you should be careful to take their signs into account. It is best to assign a reference direction to each unknown quantity, if it is not already given. The total voltage or current is calculated as the algebraic sum of the contributions from the sources. If a contribution from a source has the same direction as the reference direction, it has a positive sign in the sum; if it has the opposite direction, then a negative sign.

Note that If the voltage or current sources have internal resistance, it must remain in the circuit and still be considered. In TINA, you can assign an internal resistance to the DC voltage and current sources, while using the same schematic symbol. Therefore, if you want to illustrate the superposition theorem and at the same time use sources with internal resistance, you should only set the source voltage (or current) to zero, which leaves the source internal resistance intact. Alternatively, you could replace the source with a resistor equal to its internal resistance.

In order to use the superposition theorem with circuit currents and voltages, all of the components must be linear; that is, for all resistive components, the current must be proportional to the applied voltage (satisfying Ohms law). Note that the superposition theorem is not applicable to power, since power is not a linear quantity. The total power delivered to a resistive component must be determined using the total current through or the total voltage across the component and cannot be determined by a simple sum of the powers produced by the sources independently.

7. Wheatstone Bridge

The Wheatstone bridge is used by mechanical and civil engineers to measure the resistances of strain gauges in experimental stress studies of machines and buildings.

Topic : Inductance And Capacitance

Topic Objective:

At the end of this topic student would be able to:

- Find the current (voltage) for a capacitance or inductance given the voltage (current) as a function of time.
- Compute the capacitances of parallel-plate capacitors.
- Compute the energies stored in capacitances or inductances.
- Describe typical physical construction of capacitors and inductors and identify parasitic effects.
- Find the voltages across mutually coupled inductances in terms of the currents.

Definition/Overview:

Inductance: In electrical circuits, any electric current i flowing produces a magnetic field and hence generates a total magnetic flux Φ acting on the circuit. This magnetic flux, due to Lenz's law tends to act to oppose changes in the flux by generating a voltage (a back EMF) that counters or tends to reduce the rate of change in the current. The ratio of the magnetic flux to the current is called the self-inductance which is usually simply referred to as the inductance of the circuit.

Capacitance: Capacitance is a measure of the amount of electric charge stored (or separated) for a given electric potential. The most common form of charge storage device is a two-plate capacitor.

Key Points:

1. Capacitance

Capacitance is a measure of the amount of electric charge stored (or separated) for a given electric potential. The most common form of charge storage device is a two-plate capacitor. If the charges on the plates are +Q and −Q, and V give the voltage difference between the plates, then the capacitance is given by

The SI unit of capacitance is the farad; 1 farad = 1 coulomb per volt.

1.1. Energy

The energy (measured in joules) stored in a capacitor is equal to the work done to charge it. Consider a capacitance C, holding a charge +q on one plate and -q on the other. Moving a small element of charge dq from one plate to the other against the potential difference V = q/C requires the work dW:

Where

W is the work measured in joules

q is the charge measured in coulombs

C is the capacitance, measured in farads

We can find the energy stored in a capacitance by integrating this equation. Starting with an uncharged capacitance (q=0) and moving charge from one plate to the other until the plates have charge +Q and -Q requires the work W:

Combining this with the above equation for the capacitance of a flat-plate capacitor, we get:

.

Where

W is the energy measured in joules

C is the capacitance, measured in farads

V is the voltage measured in volts

1.2. Capacitance and 'Displacement Current'

The physicist James Clerk Maxwell invented the concept of displacement current,

, to make Ampre's law consistent with conservation of charge in cases where charge is accumulating, for example in a capacitor. He interpreted this as a real motion of charges, even in vacuum, where he supposed that it corresponded to motion of dipole charges in the ether. Although this interpretation has been abandoned, Maxwell's correction to Ampre's law remains valid (a changing electric field produces a magnetic field).

Maxwell's equation combining Ampre's law with the displacement current concept is given as . (Integrating both sides, the integral of can be replaced courtesy of Stokes's theorem with the integral of over a closed contour, thus demonstrating the interconnection with Ampre's formulation.)

1.3. Coefficients of Potential

The discussion above is limited to the case of two conducting plates, although of arbitrary size and shape. The definition C=Q/V still holds if only one plate is given a charge, provided that we recognize that the field lines produced by that charge terminate as if the plate were at the center of an oppositely charged sphere at infinity.

C=Q/V does not apply when there are more than two charged plates, or when the net charge on the two plates is non-zero. To handle this case, Maxwell introduced his "coefficients of potential". If three plates are given charges Q1, Q2, Q3, then the voltage of plate 1 is given by V1 = p11Q1 + p12Q2 + p13Q3, and similarly for the other voltages. Maxwell showed that the coefficients of potential are symmetric, so that p12 = p21, etc.

1.4. Capacitance/Inductance Duality

In mathematical terms, the ideal capacitance can be considered as an inverse of the ideal inductance, because the voltage-current equations of the two phenomena can be transformed into one another by exchanging the voltage and current terms.

1.4.1. Self-Capacitance

In electrical circuits, the term capacitance is usually a shorthand for the mutual capacitance between two adjacent conductors, such as the two plates of a capacitor. There also exists a property called self-capacitance, which is the amount of electrical charge that must be added to an isolated conductor to raise its electrical potential by one volt. The reference point for this potential is a theoretical hollow conducting sphere, of infinite radius, centred on the conductor. Using this method, the self-capacitance of a conducting sphere of radius R is given by:

Typical values of self-capacitance are:

for the top "plate" of a van de Graaf generator, typically a sphere 20 cm in radius: 20 pF

the planet Earth: about 710 F

1.4.2. Elastance

The inverse of capacitance is called elastance, and its unit is the reciprocal farad.

1.4.3. Stray Capacitance

Any two adjacent conductors can be considered as a capacitor, although the capacitance will be small unless the conductors are close together or long. This (unwanted) effect is termed "stray capacitance". Stray capacitance can allow signals to leak between otherwise isolated circuits (an effect called crosstalk), and it can be a limiting factor for proper functioning of circuits at high frequency.

Stray capacitance is often encountered in amplifier circuits in the form of "feedthrough" capacitance that interconnects the input and output nodes (both defined relative to a common ground). It is often convenient for analytical purposes to replace this capacitance with a combination of one input-to-ground capacitance and one output-to-ground capacitance. (The original configuration including the input-to-output capacitance is often referred to as a pi-configuration.) Miller's theorem can be used to effect this replacement. Miller's theorem states that, if the gain ratio of two nodes is 1:K, then an impedance of Z connecting the two nodes can be replaced with a Z/(1-k) impedance between the first node and ground and a KZ/(K-1) impedance between the second node and ground. (Since impedance varies inversely with capacitance, the internode capacitance, C, will be seen to have been replaced by a capacitance of KC from input to ground and a capacitance of (K-1)C/K from output to ground.) When the input-to-output gain is very large, the equivalent input-to-ground impedance is very small while the output-to-ground impedance is essentially equal to the original (input-to-output) impedance.

1.4.4. Capacitors

The capacitance of the majority of capacitors used in electronic circuits is several orders of magnitude smaller than the farad. The most common subunits of capacitance in use today are the millifarad (mF), microfarad (F), the nanofarad (nF) and the picofarad (pF)

The capacitance can be calculated if the geometry of the conductors and the dielectric properties of the insulator between the conductors are known. For example, the capacitance of a parallel-plate capacitor constructed of two parallel plates of area A separated by a distance d is approximately equal to the following:

(in SI units)

Where

C is the capacitance in farads, F

A is the area of each plate, measured in square metres

εr is the relative static permittivity (sometimes called the dielectric constant) of the material between the plates, (vacuum =1)

ε0 is the permittivity of free space where ε0 = 8.854x10-12 F/m

d is the separation between the plates, measured in metres

The equation is a good approximation if d is small compared to the other dimensions of the plates. In CGS units the equation has the form:

Where

C in this case has the units of length.

The dielectric constant for a number of very useful dielectrics changes as a function of the applied electrical field, e.g. ferroelectric materials, so the capacitance for these devices is no longer purely a function of device geometry. If a capacitor is driven with a sinusoidal voltage, the dielectric constant, or more accurately referred to as the relative static permittivity, is a function of frequency. A changing dielectric constant with frequency is referred to as a dielectric dispersion, and is governed by dielectric relaxation processes, such as Debye relaxation.

2. Inductance

An electric current i flowing around a circuit produces a magnetic field and hence a magnetic flux Φ through the circuit. The ratio of the magnetic flux to the current is called the inductance, or more accurately self-inductance of the circuit. The term was coined by Oliver Heaviside in February 1886. It is customary to use the symbol L for inductance, possibly in honour of the physicist Heinrich Lenz. The quantitative definition of the inductance in SI units (webers per ampere) is

In honor of Joseph Henry, the unit of inductance has been given the name Henry (H): 1H = 1Wb/A.

In the above definition, the magnetic flux Φ is that caused by the current flowing through the circuit concerned. There may, however, be contributions from other circuits. Consider for example two circuits C1, C2, carrying the currents i1, i2. The magnetic fluxes Φ1 and Φ2 in C1 and C2, respectively, are given by

According to the above definition, L11 and L22 are the self-inductances of C1 and C2, respectively. It can be shown that the other two coefficients are equal: L12 = L21 = M, where M is called the mutual inductance of the pair of circuits.

Self and mutual inductances also occur in the expression

for the energy of the magnetic field generated by N electrical circuits carrying the currents in. This equation is an alternative definition of inductance, also valid when the currents don't flow in thin wires and when it thus is not immediately clear what the area encompassed by a circuit is and how the magnetic flux through the circuit is to be defined. The definition L = Φ / i, in contrast, is more direct and more intuitive. It may be shown that the two definitions are equivalent by equating the time derivate of W and the electric power transferred to the system.

2.1. Coupled Inductors

Mutual inductance is the concept that the change in current in one inductor can induce a voltage in another nearby inductor. It is important as the mechanism by which transformers work, but it can also cause unwanted coupling between conductors in a circuit. The mutual inductance, M, is also a measure of the coupling between two inductors. The mutual inductance by circuit i on circuit j is given by the double integral Neumann formula, see #Calculation techniques

The mutual inductance also has the relationship:

where

M21 is the mutual inductance, and the subscript specifies the relationship of the voltage induced in coil 2 to the current in coil 1.

N1 is the number of turns in coil 1,

N2 is the number of turns in coil 2,

P21 is the permeance of the space occupied by the flux.

The mutual inductance also has a relationship with the coupling coefficient. The coupling coefficient is always between 1 and 0, and is a convenient way to specify the relationship between a certain orientation of inductor with arbitrary inductance:

where

k is the coupling coefficient and $0 \le k \le 1$,

L1 is the inductance of the first coil, and

L2 is the inductance of the second coil.

Once this mutual inductance factor M is determined, it can be used to predict the behavior of a circuit:

where

V is the voltage across the inductor of interest,

L1 is the inductance of the inductor of interest,

dI1 / dt is the derivative, with respect to time, of the current through the inductor of interest,

M is the mutual inductance and

dI2 / dt is the derivative, with respect to time, of the current through the inductor that is coupled to the first inductor.

When one inductor is closely coupled to another inductor through mutual inductance, such as in a transformer, the voltages, currents, and number of turns can be related in the following way:

where

Vs is the voltage across the secondary inductor,

Vp is the voltage across the primary inductor (the one connected to a power source),

Ns is the number of turns in the secondary inductor, and

Np is the number of turns in the primary inductor.

Conversely the current:

where

Is is the current through the secondary inductor,

Ip is the current through the primary inductor (the one connected to a power source),

Ns is the number of turns in the secondary inductor, and

Np is the number of turns in the primary inductor.

When either side of the transformer is a tuned circuit, the amount of mutual inductance between the two windings determines the shape of the frequency response curve. Although no boundaries are defined, this is often referred to as loose-, critical-, and over-coupling. When two tuned circuits are loosely coupled through mutual inductance, the bandwidth will be narrow. As the amount of mutual inductance increases, the bandwidth continues to grow. When the mutual inductance is increased beyond a critical point, the peak in the response curve begins to drop, and the center frequency will be attenuated more strongly than its direct sidebands. This is known as overcoupling.

Topic : Transients

Topic Objective:

At the end of this topic student would be able to:

- Solve first-order RC or RL circuits.
- Understand the concepts of transient response and steady-state response.
- Relate the transient response of first-order circuits to the time constant.
- Solve RLC circuits in dc steady-state conditions.
- Solve second-order circuits.

- Relate the step response of a second-order system to its natural frequency and damping ratio.

Definition/Overview:

Transients: The time-varying currents and voltages resulting from the sudden application of sources, usually due to switching, are called transients. By writing circuit equations, we obtain integrodifferential equations.

Key Points:

1. First Order Circuits

First order circuits are circuits that contain only one energy storage element (capacitor or inductor), and that can therefore be described using only a first order differential equation. The two possible types of first-order circuits are:

- RC (resistor and capacitor)
- RL (resistor and inductor)

RL and RC circuits is a term we will be using to describe a circuit that has either

- Resistors or inductors (RL), or
- Resistors and capacitors (RC).

These circuits are known as "First Order" circuits, because the solution to the circuit can be written as a first-order differential equation.

1.1. RL Circuits

An RL Circuit has at least one resistor (R) and one inductor (L). These can be arranged in parallel, or in series. Inductors are best solved by considering the current flowing through the inductor. Therefore, we will combine the resistive element and the source into a Norton Source Circuit. The Inductor then, will be the external load to the circuit. We remember the equation for the inductor:

If we apply KCL on the node that forms the positive terminal of the voltage source, we can solve to get the following differential equation:

1.2. RC Circuits

No, RC does not stand for "Remote Control". An RC circuit is a circuit that has both a resistor (R) and a capacitor (C). Like the RL Circuit, we will combine the resistor and the source on one side of the circuit, and combine them into a thevenin source. Then if we apply KVL around the resulting loop, we get the following equation:

2. First Order Solution

RL and RC circuits will both produce a first-order differential equation. The reader does not, however, require a prior knowledge of differential equations to read this topic, because we work through to the general solution of the equation. To understand the material fully, you would need knowledge of derivatives and integrals. We will replace the capacitor voltages and the inductor currents in the previous equations with an x to signify that this will be a general solution to either type of problem. Here, we will consider a general equation of the form:

Where k is a constant value that corresponds to the source value (current for RL and voltage for RC circuits), possibly scaled by a certain factor based on the resistance, inductance, and/or capacitance of the circuit, when we divide through. Tc is a value known as the "Time Constant".

If we separate out the variables, we can get all the x terms on one side of the equation, and all the t terms on the other:

We can integrate both sides of this equation. The left side can be integrated with respect to x, and the left side can be integrated with respect to t. Performing the integrations gives us the following equation:

Where D is an arbitrary constant of integration. If we raise both sides to e (to get rid of the natural log function) we will get the following final result:

A = eD

It turns out that A is also the value of the initial condition of the circuit, x(0). Also kTc is equal to the value of the steady-state value of the function. Combining this knowledge, we get the following equation:

[First Order Solution]

Where:

Is the steady state value of x. This is our general result. Remember that x gets replaced by the function for either the capacitor voltage or the inductor current, to get the solution to an RC or an RL circuit, respectively.

2.1. Time Constant

The Time Constant, Tc, is an indicator of the amount of time it takes for a system to react to an input. The Time Constant is based on the amount of total resistance, capacitance, and inductance of a circuit. In general, the Time constant for an RL circuit is:

and the time constant for an RC circuit is:

Tc = RC

In general, from an engineering standpoint, we say that the system is at steady state after a time period of five time constants.

3. DC Steady State

In electronics, a steady state (S.S.) occurs in a circuit or network when all transients have died away. It is an equilibrium condition that occurs as the effects of transients are no longer important. DC steady state is that state, if it exists, of a circuit described by a differential equation where the transient(s) have decayed away--typically the solution as time goes to infinity. Steady state determination is an important topic today, because many design specifications in a power electronic system are given in terms of the systems steady-state characteristics. Periodic steady-state solution is also a prerequisite for small signal dynamic modeling. Steady-state analysis is therefore an indispensable component of the design process.

In some types of circuits, such as lightly damped systems this integration could extend over many periods making the computation costly. The same happens for the so-called "stiff" circuits, and for circuits were a high frequency carrier is modulated by a much slower signal (e.g. Cellphone mixers): such circuits, for example, require a computational-costly integration over a wide span of time, with much smaller integration step. Faster numerical methods than the classical brute force integration are available to find the steady state (periodic, quasiperiodic) of non-autonomous and autonomous circuits (such as, the period T is not known a priori). Such methods are often referred as fast steady state algorithms.

3.1. Types

Steady state algorithms can be sorted into:

TD Time domain algorithms (Time domain sensitivities, Shooting)

FD Frequency domain algorithms (Harmonic Balance)

Harmonic balance methods, are the best choice for most microwave circuits excited with sinusoidal signals (e.g. mixers, power amplifiers).

3.2. Time Domain Methods

Time domain methods can be further divided into:

One step methods (Time domain Sensitivities)

Iterative methods (Shooting methods).

One step methods require Derivatives to compute the S.S; whenever those are not readily available at hand or at the output of the simulator involved, iterative methods come into focus. SPICE, for example, doesn't output derivatives, and it's not readily suitable to be the simulator of choice to compute SS thru time domain sensitivities. There's the slight option to rebuild derivatives numerically, but iterative methods are often preferred.

3.3. Steady-State Response

A steady-state response is the electrical response of a system at equilibrium. The steady-state response does not necessarily mean the response is a fixed value. An AC power supply has no fixed voltage on the output but the output is steady (a voltage of a fixed frequency and amplitude). The steady-state response follows the transient response. It is also sometimes referred to as the forced response in systems involving damping, though this is not an entirely accurate description. The forced response of a system has both transient and steady-state components.

In aerospace engineering the steady-state response is used in conjunction with the theory governing aircraft flight controls as well as for topics such as aeroacoustics, vibrations, and nonlinear dynamics. In flight control theory, for example, the steady-state response of a system can be thought of in terms of deflections of the aerodynamic control surfaces such as ailerons, elevators, and rudders. For example, if a pilot were to pull back on the yoke to pitch the aircraft up (by deflecting the elevators up) the response would be a plot of the aircraft's pitch as a function of time. This plot typically looks sinusoidal damping to a constant value. This constant value is the steady-state response -- it is the response of the aircraft after a long period of time has passed (mathematically at infinity) such that the transient response (the initial oscillations after the input) no longer has an effect.

4. RC and RL Circuits with General Sources

The general solution consists of two parts. The particular solution (also called the forced response) is any expression that satisfies the equation. In order to have a solution that satisfies the initial conditions, we must add the complementary solution to the particular solution. The homogeneous equation is obtained by setting the forcing function to zero. The complementary solution (also called the natural response) is obtained by solving the homogeneous equation.

4.1. Step-by-Step Solution

Circuits containing a resistance, a source, and an inductance (or a capacitance)

Write the circuit equation and reduce it to a first-order differential equation.

Find a particular solution. The details of this step depend on the form of the forcing function. We illustrate several types of forcing functions in examples, exercises, and problems.

Obtain the complete solution by adding the particular solution to the complementary solution given by Equation 4.44, which contains the arbitrary constant K.

Use initial conditions to find the value of K.

In Section 2 of this course you will cover these topics:

- Steady-State Sinusoidal Analysis
- Frequency Response, Bode Plots, And Resonance
- Logic Circuits
- Microcomputers

Topic : Steady-State Sinusoidal Analysis

Topic Objective:

At the end of this topic student would be able to:

- Identify the frequency, angular frequency, peak value, rms value, and phase of a sinusoidal signal.
- Determine the rms value of any periodic current or voltage.
- Solve steady-state ac circuits, using phasors and complex impedances.
- Compute power for steady-state ac circuits.
- Find Thvenin and Norton equivalent circuits.
- Determine load impedances for maximum power transfer.

- Discuss the advantages of three-phase power distribution.
- Solve balanced three-phase circuits.

Definition/Overview:

Frequency: Frequency is a measure of the number of occurrences of a repeating event per unit time. It is also referred to as temporal frequency. The period is the duration of one cycle in a repeating event, so the period is the reciprocal of the frequency.

Angular Frequency: In physics (specifically mechanics and electrical engineering), angular frequency ω is a scalar measure of rotation rate. Angular frequency (or angular speed) is the magnitude of the vector quantity angular velocity.

Sinusoidal Signal: Sinusoidal signals are often voltages which vary sinusoidally in time. (Sinusoidal signals could be, however, other physical variables like current, pressure, or virtually any other physical variable.) Here's a simulator that will let you put various kinds of signals on a simulated oscilloscope.

Key Points:

1. RMS Value of a Sinusoid

The rms value for a sinusoid is the peak value divided by the square root of two. This is not true for other periodic waveforms such as square waves or triangular waves.

2. Phasor

In physics and engineering, a phasor is a representation of a sine wave whose amplitude (A), phase (θ), and frequency (ω) are time-invariant. It is a subset of a more general concept called analytic representation. Phasors reduce the dependencies on these parameters to three independent factors, thereby simplifying certain kinds of calculations. In particular, the frequency factor, which also includes the time-dependence of the sine wave, is often common to all the components of a linear combination of sine waves. Using phasors, it can be factored out, leaving just the static amplitude and phase information to be combined algebraically (rather than trigonometrically). Similarly, linear differential equations can be reduced to algebraic ones. The term phasor therefore often refers to just those two factors. In older texts, a phasor is also referred to as a sinor.

Euler's formula indicates that sine waves can be represented mathematically as the sum of two complex-valued functions:

or as the real part of one of the functions:

As indicated above, phasor can refer to either or just the complex constant, . In the latter case, it is understood to be a shorthand notation, encoding the amplitude and phase of an underlying sinusoid.

Even more compact shorthand is angle notation:

2.1. Adding Sinusoids Using Phasors

Step 1: Determine the phasor for each term.

Step 2: Add the phasors using complex arithmetic.

Step 3: Convert the sum to polar form.

Step 4: Write the result as a time function.

3. Phase Relationships

To determine phase relationships from a phasor diagram, consider the phasors to rotate counterclockwise. Then when standing at a fixed point, if V1 arrives first followed by V2 after a rotation of θ , we say that V1 leads V2 by θ . Alternatively, we could say that V2 lags V1 by θ . (Usually, we take θ as the smaller angle between the two phasors.)

To determine phase relationships between sinusoids from their plots versus time, find the shortest time interval tp between positive peaks of the two waveforms. Then, the phase angle is θ = (tp/T) 360. If the peak of v1(t) occurs first, we say that v1(t) leads v2(t) or that v2(t) lags v1(t).

4. Complex Impedance

The handling of the impedance of an AC circuit with multiple components quickly becomes unmanageable if sines and cosines are used to represent the voltages and currents. A mathematical construct which eases the difficulty is the use of complex exponential functions.

5. Kirchhoffs Laws in Phasor Form

We can apply KVL directly to phasors. The sum of the phasor voltages equals zero for any closed path. The sum of the phasor currents entering a node must equal the sum of the phasor currents leaving.

6. Circuit Analysis Using Phasors and Impedances

- Replace the time descriptions of the voltage and current sources with the corresponding phasors. (All of the sources must have the same frequency.)
- Replace inductances by their complex impedances ZL = jωL. Replace capacitances by their complex impedances ZC = 1/(jωC). Resistances have impedances equal to their resistances.
- Analyze the circuit using any of the techniques studied earlier in Chapter 2, performing the calculations with complex arithmetic.

7. Thvenin Equivalent Circuits

The Thvenin voltage is equal to the open-circuit phasor voltage of the original circuit.

We can find the Thvenin impedance by zeroing the independent sources and determining the impedance looking into the circuit terminals.

8. Maximum Power Transfer

- If the load can take on any complex value, maximum power transfer is attained for a load impedance equal to the complex conjugate of the Thvenin impedance.
- If the load is required to be a pure resistance, maximum power transfer is attained for a load resistance equal to the magnitude of the Thvenin impedance.

Topic : Frequency Response, Bode Plots, And Resonance

Topic Objective:

At the end of this topic student would be able to:

- State the fundamental concepts of Fourier analysis.
- Use a filters transfer function to determine its output for a given input consisting of sinusoidal components.
- Use circuit analysis to determine the transfer functions of simple circuits.
- Draw first-order lowpass or highpass filter circuits and sketch their transfer functions.
- Understand decibels, logarithmic frequency scales, and Bode plots.
- Draw the Bode plots for transfer functions of firstorder filters.
- Use software to produce Bode plots for more complex RLC filters.
- Calculate parameters for series- and parallelresonant circuits.
- Select and design simple filter circuits.
- Design simple digital signal-processing systems.

Definition/Overview:

Fourier Analysis: All real-world signals are sums of sinusoidal components having various frequencies, amplitudes, and phases.

Filters: Filters process the sinusoid components of an input signal differently depending of the frequency of each component. Often, the goal of the filter is to retain the components in certain frequency ranges and to reject components in other ranges.

Key Points:

1. Fourier Analysis

In mathematics, Fourier analysis is a subject area which grew out of the study of Fourier series. The subject began with trying to understand when it was possible to represent general functions by sums of simpler trigonometric functions. The attempt to understand functions (or other objects) by breaking them into basic pieces that are easier to understand is one of the central themes in Fourier analysis.

Today the subject of Fourier analysis encompasses a vast spectrum of mathematics with parts that, at first glance, may appear quite different. In the sciences and engineering the process of decomposing a function into simpler pieces is often called an analysis. The corresponding operation of rebuilding the function from these pieces is known as synthesis. In this context the term Fourier synthesis describes the act of rebuilding and the term Fourier analysis describes the process of breaking the function into a sum of simpler pieces. In mathematics, the term Fourier analysis often refers to the study of both operations.

In Fourier analysis, the term Fourier transform often refers to the process that decomposes a given function into the basic pieces. This process results in another function that describes how much of each basic piece are in the original function. It is common practice to also use the term Fourier transform to refer to this function. However, the transform is often given a more specific name depending upon the domain and other properties of the function being transformed, as elaborated below. Moreover, the original concept of Fourier analysis has been extended over time to apply to more and more abstract and general situations, and the general field is often known as harmonic analysis.

2. Filters

2.1. Transfer Functions

The transfer function H(f) of the two-port filter is defined to be the ratio of the phasor output voltage to the phasor input voltage as a function of frequency:

The magnitude of the transfer function shows how the amplitude of each frequency component is affected by the filter. Similarly, the phase of the transfer function shows how the phase of each frequency component is affected by the filter.

2.2. Determining the Output

Determining the output of a filter for an input with multiple components:

- Determine the frequency and phasor representation for each input component.
- Determine the (complex) value of the transfer function for each component.
- Obtain the phasor for each output component by multiplying the phasor for each input component by the corresponding transfer-function value.
- Convert the phasors for the output components into time functions of various frequencies. Add these time functions to produce the output.

2.3. Linear Circuits Behaviour

Linear circuits behave as if they:

- Separate the input signal into components having various frequencies.
- Alter the amplitude and phase of each component depending on its frequency.
- Add the altered components to produce the output signal.

3. Low-Pass Filter

A low-pass filter is a filter that passes low-frequency signals but attenuates (reduces the amplitude of) signals with frequencies higher than the cutoff frequency. The actual amount of attenuation for each frequency varies from filter to filter. It is sometimes called a high-cut filter, or treble cut filter when used in audio applications.

The concept of a low-pass filter exists in many different forms, including electronic circuits (like a hiss filter used in audio), digital algorithms for smoothing sets of data, acoustic barriers, blurring of images, and so on. Low-pass filters play the same role in signal processing that moving averages do in some other fields, such as finance; both tools provide a smoother form of a signal which removes the short-term oscillations, leaving only the long-term trend.

3.1. Examples of Low Pass Filters

Figure 1 shows a low pass RC filter for voltage signals, discussed in more detail below. Signal Vout contains frequencies from the input signal, with high frequencies attenuated, but with little attentuation below the corner frequency of the filter determined by its RC time constant. For current signals, a similar circuit using a resistor and capacitor in parallel works the same way.

3.2. Acoustic

A stiff physical barrier tends to reflect higher sound frequencies, and so acts as a low-pass filter for transmitting sound. When music is playing in another room, the low notes are easily heard, while the high notes are attenuated.

3.3. Electronic

- Electronic low-pass filters are used to drive subwoofers and other types of loudspeakers, to block high pitches that they can't efficiently broadcast.
- Radio transmitters use lowpass filters to block harmonic emissions which might cause interference with other communications.

- An integrator is another example of a low-pass filter.
- DSL splitters use low-pass and high-pass filters to separate DSL and POTS signals sharing the same pair of wires.
- Low-pass filters also play a significant role in the sculpting of sound for electronic music as created by analogue synthesisers.

3.4. Ideal and Real Filters

An ideal low-pass filter completely eliminates all frequencies above the cut-off frequency while passing those below unchanged. The transition region present in practical filters does not exist in an ideal filter. An ideal low-pass filter can be realized mathematically (theoretically) by multiplying a signal by the rectangular function in the frequency domain or, equivalently, convolution with a sinc function in the time domain.

However, the ideal filter is impossible to realize without also having signals of infinite extent, and so generally needs to be approximated for real ongoing signals, because the sinc function's support region extends to all past and future times. The filter would therefore need to have infinite delay, or knowledge of the infinite future and past, in order to perform the convolution. It is effectively realizable for pre-recorded digital signals by assuming extensions of zero into the past and future, but even that is not typically practical.

Real filters for real-time applications approximate the ideal filter by truncating and windowing the infinite impulse response to make a finite impulse response; applying that filter requires delaying the signal for a moderate period of time, allowing the computation to "see" a little bit into the future. This delay is manifested as phase shift. Greater accuracy in approximation requires a longer delay.

The WhittakerShannon interpolation formula describes how to use a perfect low-pass filter to reconstruct a continuous signal from a sampled digital signal. Real digital-to-analog converters use real filter approximations.

3.5. Electronic Low-Pass Filters

There are a great many different types of filter circuits, with different responses to changing frequency. The frequency response of a filter is generally represented using a Bode plot. A first-order filter, for example, will reduce the signal amplitude by half (about 6 dB) every time the frequency doubles (goes up one octave). The magnitude Bode plot for a first-order filter looks like a horizontal line below the cutoff frequency, and a diagonal line above the cutoff frequency. There is also a "knee curve" at the boundary between the two, which smoothly transitions between the two straight line regions.

A second-order filter does a better job of attenuating higher frequencies. The Bode plot for this type of filter resembles that of a first-order filter, except that it falls off more quickly. For example, a second-order Butterworth filter will reduce the signal amplitude to one fourth its original level every time the frequency doubles (12 dB per octave). Other second-order filters may roll off at different rates initially depending on their Q factor, but approach the same final rate of 12 dB per octave.

Third- and higher-order filters are defined similarly. In general, the final rate of rolloff for an n-order filter is 6n dB per octave. On any Butterworth filter, if one extends the horizontal line to the right and the diagonal line to the upper-left (the asymptotes of the function), they will intersect at exactly the "cutoff frequency". The frequency response at the cutoff frequency in a first-order filter is 3 dB below the horizontal line. The various types of filters Butterworth filter, Chebyshev filter, Bessel filter, etc. all have different-looking "knee curves". Many second-order filters are designed to have "peaking" or resonance, causing their frequency response at the cutoff frequency to be above the horizontal line. The meanings of 'low' and 'high' that is, the cutoff frequency depend on the characteristics of the filter. The term "low-pass filter" merely refers to the shape of the filter's response; a high-pass filter could be built that cuts off at a lower frequency than any low-pass filter it is their responses that set them apart. Electronic circuits can be devised for any desired frequency range, right up through microwave frequencies (above 1000 MHz) and higher.

3.6. Passive Electronic Realization

One simple electrical circuit that will serve as a low-pass filter consists of a resistor in series with a load, and a capacitor in parallel with the load. The capacitor exhibits reactance, and blocks low-frequency signals, causing them to go through the load instead. At higher frequencies the reactance drops, and the capacitor effectively functions as a short circuit. The combination of resistance and capacitance gives you the time constant of the filter $\tau = RC$ (represented by the Greek letter tau). The break frequency, also called the turnover frequency or cutoff frequency (in hertz), is determined by the time constant:

or equivalently (in radians per second):

One way to understand this circuit is to focus on the time the capacitor takes to charge. It takes time to charge or discharge the capacitor through that resistor:

At low frequencies, there is plenty of time for the capacitor to charge up to practically the same voltage as the input voltage. At high frequencies, the capacitor only has time to charge up a small amount before the input switches direction. The output goes up and down only a small fraction of the amount the input goes up and down. At double the frequency, there's only time for it to charge up half the amount. Another way to understand this circuit is with the idea of reactance at a particular frequency:

- Since DC cannot flow through the capacitor, DC input must "flow out" the path marked Vout (analogous to removing the capacitor).
- Since AC flows very well through the capacitor almost as well as it flows through solid wire AC input "flows out" through the capacitor, effectively short circuiting to ground (analogous to replacing the capacitor with just a wire).
- It should be noted that the capacitor is not an "on/off" object (like the block or pass fluidic explanation above). The capacitor will variably act between these two extremes. It is the Bode plot and frequency response that show this variability.

3.7. Active Electronic Realization

Another type of electrical circuit is an active low-pass filter.

In the operational amplifier circuit shown in the figure, the cutoff frequency (in hertz) is defined as:

or equivalently (in radians per second):

The gain in the passband is , and the stopband drops off at −6 dB per octave, as it is a first-order filter.

Sometimes, a simple gain amplifier (as opposed to the very-high-gain operation amplifier) is turned into a low-pass filter by simply adding a feedback capacitor C. This feedback decreases the frequency response at high frequencies via the Miller effect, and helps to avoid oscillation in the amplifier. For example, an audio amplifier can be made into a low-pass filter with cutoff frequency 100 kHz to reduce gain at frequencies which would otherwise oscillate. Since the audio band (what we can hear) only goes up to 20 kHz or so, the frequencies of interest fall entirely in the passband, and the amplifier behaves the same way as far as audio is concerned.

4. Decibel

The decibel (dB) is a logarithmic unit of measurement that expresses the magnitude of a physical quantity (usually power or intensity) relative to a specified or implied reference level. Since it expresses a ratio of two quantities with the same unit, it is a dimensionless unit. A decibel is one tenth of a bel (B). The decibel is useful for a wide variety of measurements in science and engineering (e.g., acoustics and electronics) and other disciplines. It confers a number of advantages, such as the ability to conveniently represent very large or small numbers, a logarithmic scaling that roughly corresponds to the human perception of, for example, sound and light, and the ability to carry out multiplication of ratios by simple addition and subtraction.

5. Bode Plot

A Bode plot, named after Hendrik Wade Bode, is usually a combination of a Bode magnitude plot and Bode phase plot.

5.1. Bode Magnitude Plot

A Bode magnitude plot is a graph of log magnitude versus frequency, plotted with a log-frequency axis, to show the transfer function or frequency response of a linear, time-invariant system. The magnitude axis of the Bode plot is usually expressed as decibels, that is, 20 times the common logarithm of the amplitude gain. With the magnitude gain being logarithmic, Bode plots make multiplication of magnitudes a simple matter of adding distances on the graph (in decibels), since

5.2. Bode Phase Plot

A Bode phase plot is a graph of phase versus frequency, also plotted on a log-frequency axis, usually used in conjunction with the magnitude plot, to evaluate how much a frequency will be phase-shifted. For example a signal described by: $A\sin(\omega t)$ may be attenuated but also phase-shifted. If the system attenuates it by a factor x and phase shifts it by $-\Phi$ the signal out of the system will be $(A/x)\sin(\omega t - \Phi)$. The phase shift Φ is generally a function of frequency. Phase can also be added directly from the graphical values, a fact that is mathematically clear when phase is seen as the imaginary part of the complex logarithm of a complex gain.

6. Resonance

Resonance in AC circuits implies a special frequency determined by the values of the resistance , capacitance , and inductance . For series resonance the condition of resonance is straightforward and it is characterized by minimum impedance and zero phase. Parallel resonance , which is more common in electronic practice, requires a more careful definition.

6.1. Series Resonance

The resonance of a series RLC circuit occurs when the inductive and capacitive reactances are equal in magnitude but cancel each other because they are 180 degrees apart in phase. The sharp minimum in impedance which occurs is useful in tuning applications. The sharpness of the minimum depends on the value of R and is characterized by the "Q" of the circuit.

6.2. Parallel Resonance

The resonance of a parallel RLC circuit is a bit more involved than the series resonance. The resonant frequency can be defined in three different ways, which converge on the same expression as the series resonant frequency if the resistance of the circuit is small.

Topic : Logic Circuits

Topic Objective:

At the end of this topic student would be able to:

- State the advantages of digital technology over analog technology.
- Understand the terminology of digital circuits.
- Convert numbers between decimal, binary, and other forms.
- Use the Gray code for position and angular sensors.
- Understand the binary arithmetic operations used in computers and other digital systems.
- Interconnect logic gates of various types to implement a given logic function.
- Use Karnaugh maps to minimize the number of gates needed to implement a logic function.
- Understand how gates are connected together to form flip-flops and registers.

Definition/Overview:

Logic Circuit: logic circuit, electric circuit whose output depends upon the input in a way that can be expressed as a function in symbolic logic; it has one or more binary inputs (capable of assuming either of two states, e.g., on or off) and a single binary output.

Truth Table: A truth table is a mathematical table used in logic specifically in connection with Boolean algebra, boolean functions, and propositional calculus to compute the functional values of logical expressions on each of their functional arguments, that is, on each combination of values taken by their logical variables. In particular, truth tables can be used to tell whether a propositional expression is true for all legitimate input values, that is, logically valid.

Key Points:

1. Logic Circuits

logic circuit, electric circuit whose output depends upon the input in a way that can be expressed as a function in symbolic logic; it has one or more binary inputs (capable of assuming either of two states, e.g., on or off) and a single binary output. Logic circuits that perform particular functions are called gates. Basic logic circuits include the AND gate, the OR gate, and the NOT gate, which perform the logical functions AND, OR, and NOT. Logic circuits can be built from any binary electric or electronic devices, including switches, relays, electron tubes, solid-state diodes, and transistors; the choice depends upon the application and design requirements. Modern technology has produced integrated logic circuits, modules that perform complex logical functions. A major use of logic circuits is in electronic digital computers. Fluid logic circuits have been developed whose function depends on the flow of a liquid or gas rather than on an electric current.

1.1. Logic Gate

A logic gate performs a logical operation on one or more logic inputs and produces a single logic output. Because the output is also a logic-level value, an output of one logic gate can connect to the input of one or more other logic gates. The logic normally performed is Boolean logic and is most commonly found in digital circuits. Logic gates are primarily implemented electronically using diodes or transistors, but can also be constructed using electromagnetic relays, fluidics, optics, molecules, or even mechanical elements.

In electronic logic, a logic level is represented by a voltage or current, (which depends on the type of electronic logic in use). Each logic gate requires power so that it can source and sink currents to achieve the correct output voltage. In logic circuit diagrams the power is not shown, but in a full electronic schematic, power connections are required.

1.2 Logic Gates and Hardware

NAND and NOR logic gates are the two pillars of logic, in that all other types of Boolean logic gates (i.e., AND, OR, NOT, XOR, XNOR) can be created from a suitable network of just NAND or just NOR gate(s). They can be built from relays or transistors, or any other technology that can create an inverter and a two-input AND or OR gate. Hence the NAND and NOR gates are called the universal gates.

For an input of 2 variables, there are 16 possible boolean algebraic functions. These 16 functions are enumerated below, together with their outputs for each combination of inputs variables.

The four functions labeled with a "*" are the logical implication functions: "A ≤ B" can be read as "A implies B"; it follows that "A > B" is "A does not imply B". These four functions are less common and are usually not implemented directly as logic gates.

1.3. Combinational Logic

In digital circuit theory, combinational logic (also called combinatorial logic) is a type of logic circuit whose output is a pure function of the present input only. This is in contrast to sequential logic, in which the output depends not only on the present input but also on the history of the input.

In other words, sequential logic has memory while combinational logic does not.

Combinational logic is used in computer circuits to do Boolean algebra on input signals and on stored data. Practical computer circuits normally contain a mixture of combinational and sequential logic. For example, the part of an arithmetic logic unit, or ALU, that does mathematical calculations is constructed in accord with combinational logic, although the ALU is controlled by a sequencer that is constructed in accord with sequential logic.

2. Digital Circuits

Digital electronics are electronics systems that use digital signals. Digital electronics are representations of Boolean algebra also see truth tables and are used in computers, mobile phones, and other consumer products. In a digital circuit, a signal is represented in one of two states or logic levels. The advantages of digital techniques stem from the fact it is easier to get an electronic device to switch into one of two states, then to accurately reproduce a continuous range of values. Digital electronics or any digital circuit are usually made from large assemblies of logic gates, simple electronic representations of Boolean logic functions.

2.1. Advantages of the Digital Approach

Provided that the noise amplitude is not too large, the logic values represented by a digital signal can still be determined after noise is added. With modern IC technology, it is possible to manufacture exceedingly complex digital circuits economically.

2.2. Positive versus Negative Logic: Digital Words

In parallel transmission, an n-bit word is transferred on n wires, one wire for each bit, plus a common or ground wire. In serial transmission, the successive bits of the word are transferred one after the other with a single pair of wires.

3. Binary Numbers

The binary numeral system, or base-2 number system, is a numeral system that represents numeric values using two symbols, usually 0 and 1. More specifically, the usual base-2 system is a positional notation with a radix of 2. Owing to its straightforward implementation in digital electronic circuitry using logic gates, the binary system is used internally by all modern computers.

3.1. Gray Code

The reflected binary code, also known as Gray code after Frank Gray, is a binary numeral system where two successive values differ in only one digit. The reflected binary code was originally designed to prevent spurious output from electromechanical switches. Today, Gray codes are widely used to facilitate error correction in digital communications such as digital terrestrial television and some cable TV systems.

Sometimes digital buses in electronic systems are used to convey quantities that can only increase or decrease by one at a time, for example the output of an event counter which is being passed between clock domains or to a digital-to-analog converter. The advantage of Gray code in these applications is that differences in the propagation delays of the many wires that represent the bits of the code cannot cause the received value to go through states that are out of the Gray code sequence. This is similar to the advantage of Gray codes in the construction of mechanical encoders, however the source of the Gray code is an electronic counter in this case. The counter itself must count in Gray code, or if the counter runs in binary then the output value from the counter must be reclocked after it has been converted to Gray code, because when a value is converted from binary to Gray code, it is possible that differences in the arrival times of the binary data bits into the binary-to-Gray conversion circuit will mean that the code could go briefly through states that are wildly out of sequence. Adding a clocked register after the circuit that converts the count value to Gray code may introduce a clock cycle of latency, so counting directly in Gray code may be advantageous.

3.2. Complement Arithmetic

Most computers use complement arithmetic for integer representations. The reason for this is mostly to simplify the circuitry required to perform integer arithmetic operations. Negative numbers may represented in complement form and that the operation of subtraction may be accomplished by adding the complement of a number. We will show that the complement of a number is very easy to calculate and both addition and subtraction can be accomplished by adding!

The ones complement of a binary number is obtained by replacing 1s by 0s, and vice versa. The twos complement of a binary number is obtained by adding 1 to the ones complement, neglecting the carry (if any) out of the most significant bit. Complements are useful for representing negative numbers and performing subtraction in computers.

4. Overflow and Underflow

In performing arithmetic using twos-complement arithmetic, we must be aware of the possibility of overflow in which the result exceeds the maximum value that can be represented by the word length in use.

5. De Morgan's Laws

In logic, De Morgan's laws or De Morgan's theorem are rules in formal logic relating pairs of dual logical operators in a systematic manner expressed in terms of negation. The relationship so induced is called De Morgan duality.

not (P and Q) = (not P) or (not Q)

not (P or Q) = (not P) and (not Q)

De Morgan's laws are based on the equivalent truth-values of each pair of statements. De Morgan's formulation was influenced by algebraisation of logic undertaken by George Boole, which later cemented De Morgan's claim to the find. Although a similar observation was made by Aristotle and was known to Greek and Medieval logicians, De Morgan is given credit for stating the laws formally and incorporating them in to the language of logic. De Morgan's Laws can be proved easily, and may even seem trivial. Nonetheless, these laws are helpful in making valid inferences in proofs and deductive arguments.

In electrical engineering contexts, the negation operator can be written as an overline above the terms to be negated, as shown above. Thus, electrical engineering students are often taught to remember De Morgan's laws using the mnemonic "break the line, change the sign".

6. Truth Table

Truth tables are used to compute the values of propositional expressions in an effective manner that is sometimes referred to as a *decision procedure*. A propositional expression is either an atomic formula a propositional constant, propositional variable, or propositional function term (for example, *Px* or *P*(*x*)) or built up from atomic formulas by means of logical operators, for example, AND ( ), OR ( ), NOT ( ). For instance, is a propositional expression. The column headings on a truth table show (i) the propositional functions and/or variables, and (ii) the truth-functional expression built up from those propositional functions or variables and operators. The rows show each possible valuation of T or F assignments to (i) and (ii). In other words, each row is a distinct interpretation of (i) and (ii). Truth tables for classical logic are limited to Boolean logical systems in which only two logical values are possible, false and true, usually written F and T, or sometimes 0 or 1, respectively.

6.1. Sum-of-Products Implementation

Product terms that include all of the input variables (or their inverses) are called minterms. In a sum-of-products expression, we form a product of all the input variables (or their inverses) for each row of the truth table for which the result is logic 1. The output is the sum of these products.

6.2. Product-of-Sums Implementation

Sum terms that include all of the input variables (or their inverses) are called maxterms. In a product-of-sums expression, we form a sum of all the input variables (or their inverses) for each row of the truth table for which the result is logic 0. The output is the product of these sums.

6.3. Karnaugh Maps

The Karnaugh map, also known as a Veitch diagram (KV-map or K-map for short), is a tool to facilitate the simplification of Boolean algebra IC expressions. The Karnaugh map reduces the need for extensive calculations by taking advantage of human pattern-recognition and permitting the rapid identification and elimination of potential race hazards.

In a Karnaugh map the boolean variables are transferred (generally from a truth table) and ordered according to the principles of Gray code in which only one variable changes in between squares. Once the table is generated and the output possibilities are transcribed, the data is arranged into the largest even group possible and the minterm is generated through the axiom laws of boolean algebra.

Topic : Microcomputers

Topic Objective:

At the end of this topic student would be able to:

- Identify and describe the functional blocks of a microcomputer.
- Select the type of memory needed for a given application.
- Understand how microcomputers ormicrocontrollers can be applied in your field of specialization.
- Identify the internal registers and their functions for the 68HC11 microcomputer.
- List some of the instructions and addressing modes of the 68HC11.
- Write simple programs, using the 68HC11 instruction set.

Definition/Overview:

Microcomputer: A microcomputer is a computer with a microprocessor as its central processing unit. Another general characteristic of these computers is that they occupy physically small amounts of space when compared to mainframe and minicomputers. Many microcomputers (when equipped with a keyboard and screen for input and output) are also personal computers (in the generic sense).

Key Points:

1. Microcomputer

An embedded computer is part of a product, such as an automobile, printer, or bread machine, that is not called a computer. A microcomputer or microcontroller is a complete computer containing the CPU, memory, and I/O on a single silicon chip.

1.1. Description

Monitors, keyboards and other devices for input and output may be integrated or separate. Computer memory in the form of RAM, and at least one other less volatile, memory storage device are usually combined with the CPU on a system bus in a single unit. Other devices that make up a complete microcomputer system include, batteries, a power supply unit, a keyboard and various input/output devices used to convey information to and from a human operator (printers, monitors, human interface devices) Microcomputers are designed to serve only a single user at a time, although they can often be modified with software or hardware to concurrently serve more than one user. Microcomputers fit well on or under desks or tables, so that they are within easy access of the user. Bigger computers like minicomputers, mainframes, and supercomputers take up large cabinets or even a dedicated room.

A microcomputer comes equipped with at least one type of data storage, usually RAM. Although some microcomputers (particularly early 8-bit home micros) perform tasks using RAM alone, some form of secondary storage is normally desirable. In the early days of home micros, this was often a data cassette deck (in many cases as an external unit). Later, secondary storage (particularly in the form of floppy disk and hard disk drives) were built in to the microcomputer case itself.

2. Memory Types

2.1. Read-and-write memory (RAM)

Random-access memory (usually known by its acronym, RAM) is a computer data storage. Today it takes the form of integrated circuits that allow the stored data to be accessed in any order, i.e. at random. The word random thus refers to the fact that any piece of data can be returned in a constant time, regardless of its physical location and whether or not it is related to the previous piece of data.

2.2. Read-only memory (ROM)

Read-only memory (usually known by its acronym, ROM) is a class of storage media used in computers and other electronic devices. Because data stored in ROM cannot be modified (at least not very quickly or easily), it is mainly used to distribute firmware (software that is very closely tied to specific hardware, and unlikely to require frequent updates).

2.3. Mass storage

In computing, mass storage refers to the storage of large amounts of information in a persisting and machine-readable fashion. Storage media for mass storage include hard disks, floppy disks, flash memory, optical discs, magneto-optical discs, magnetic tape, drum memory, punched tape (mostly historic) and holographic memory (experimental). Mass storage includes devices with removable and non-removable media. It does not include random access memory (RAM), which is volatile in that it loses its contents after power loss.

2.4. Selection of Memory

- o The trade-off between speed and cost
- o Whether the information is to be stored permanently or must be changed frequently
- o Whether data are to be accessed in random order or in sequence

In Section 3 of this course you will cover these topics:

- Computer-Based Instrumentation Systems
- Diodes
- Amplifiers: Specifications And External Characteristics

Topic : Computer-Based Instrumentation Systems

Topic Objective:

At the end of this topic student would be able to:

- Describe the operation of the elements of a computer-based instrumentation system.
- Identify the types of errors that may be encountered in instrumentation systems.
- Avoid common pitfalls such as ground loops, noise coupling, and loading when using sensors.
- Determine specifications for the elements of computer- based instrumentation systems such as dataacquisition boards.

- Know how to use LabVIEW to create virtual instruments for computer-aided test and control systems in your field of engineering.

Definition/Overview:

Instrumentation Systems: An instrument in a measurement system is a device designed specifically for collecting data from an environment or from a unit under test and to display information based on the collected data. The instruments used in measuremens are known as measurement system. The general Measurement system follows a simple block diagram. A traditional measurement system or traditional instrumentation system involves sensors or transducers, signal conditioning elements, signal processing elements, display devices. A sensor senses changes in a physical parameter such as temperature, pressure,level etc, and converted into electrical signal (voltage, current,frequency etc.) through signal conditioning elements (such as deflection bridges, amplifiers, current transmitters etc.) according to the requirements of the user. Now these electrical signal are processed under signal processing elements for further operations such as display or data storage.

Key Points:

1. Sensors

A sensor is a device that measures a physical quantity and converts it into a signal which can be read by an observer or by an instrument. For example, a mercury thermometer converts the measured temperature into expansion and contraction of a liquid which can be read on a calibrated glass tube. A thermocouple converts temperature to an output voltage which can be read by a voltmeter. For accuracy, all sensors need to be calibrated against known standards.

Sensors are used in everyday objects such as touch-sensitive elevator buttons and lamps which dim or brighten by touching the base. There are also innumerable applications for sensors of which most people are never aware. Applications include automobiles, machines, aerospace, medicine, industry, and robotics.

A sensor's sensitivity indicates how much the sensor's output changes when the measured quantity changes. For instance, if the mercury in a thermometer moves 1cm when the temperature changes by 1, the sensitivity is 1cm/1. Sensors that measure very small changes must have very high sensitivities.

Technological progress allows more and more sensors to be manufactured on a microscopic scale as microsensors using MEMS technology. In most cases, a microsensor reaches a significantly higher speed and sensitivity compared with macroscopic approaches.

2. Classification of Measurement Errors

A good sensor obeys the following rules:

- The sensor should be sensitive to the measured property
- The sensor should be insensitive to any other property
- The sensor should not influence the measured property

Ideal sensors are designed to be linear. The output signal of such a sensor is linearly proportional to the value of the measured property. The sensitivity is then defined as the ratio between output signal and measured property. For example, if a sensor measures temperature and has a voltage output, the sensitivity is a constant with the unit [V/K]; this sensor is linear because the ratio is constant at all points of measurement.

If the sensor is not ideal, several types of deviations can be observed:

- The sensitivity may in practice differ from the value specified. This is called a sensitivity error, but the sensor is still linear.

- Since the range of the output signal is always limited, the output signal will eventually reach a minimum or maximum when the measured property exceeds the limits. The full scale range defines the maximum and minimum values of the measured property.
- If the output signal is not zero when the measured property is zero, the sensor has an offset or bias. This is defined as the output of the sensor at zero input.
- If the sensitivity is not constant over the range of the sensor, this is called nonlinearity. Usually this is defined by the amount the output differs from ideal behavior over the full range of the sensor, often noted as a percentage of the full range.
- If the deviation is caused by a rapid change of the measured property over time, there is a dynamic error. Often, this behaviour is described with a bode plot showing sensitivity error and phase shift as function of the frequency of a periodic input signal.
- If the output signal slowly changes independent of the measured property, this is defined as drift.
- Long term drift usually indicates a slow degradation of sensor properties over a long period of time.
- Noise is a random deviation of the signal that varies in time.
- Hysteresis is an error caused by when the measured property reverses direction, but there is some finite lag in time for the sensor to respond, creating a different offset error in one direction than in the other.
- If the sensor has a digital output, the output is essentially an approximation of the measured property. The approximation error is also called digitization error.
- If the signal is monitored digitally, limitation of the sampling frequency also can cause a dynamic error.
- The sensor may to some extent be sensitive to properties other than the property being measured. For example, most sensors are influenced by the temperature of their environment.

All these deviations can be classified as systematic errors or random errors. Systematic errors can sometimes be compensated for by means of some kind of calibration strategy. Noise is a random error that can be reduced by signal processing, such as filtering, usually at the expense of the dynamic behaviour of the sensor.

3. Measurement Systems

- Accuracy: The maximum expected difference in magnitude between measured and true values (often expressed as a percentage of the full-scale value).
- Precision: The ability of the instrument to repeat the measurement of a constant measurand. More precise measurements have less random error.
- Resolution: The smallest possible increment discernible between measured values. As the term is used, higher resolution means smaller increments. Thus, an instrument with a five-digit display (say, 0.0000 to 9.9999) is said to have higher resolution than an otherwise identical instrument with a three-digit display (say, 0.00 to 9.99).

4. Signal Conditioning

Some functions of signal conditioners are:

- Amplification of the sensor signals
- Conversion of currents to voltages
- Supply of (ac or dc) excitations to the sensors so changes in resistance, inductance, or capacitance are converted to changes in voltage
- Filtering to eliminate noise or other unwanted signal components

5. Analog-To-Digital Conversion

If a signal contains no components with frequencies higher than fH , all of the information contained in the signal is present in its samples, provided that the sampling rate is selected to be more than twice fH . Analog-to-digital conversion is a two-step process. First, the signal is sampled at uniformly spaced points in time. Second, the sample values are quantized so they can be represented by words of finite length.

6. LabVIEW

LabVIEW, a product of National Instruments, is an industry-standard program used by all types of engineers and scientists for developing sophisticated instrumentation systems such as the timefrequency vibration analyzer. LabVIEW is an acronym for Laboratory Virtual Instrument Engineering Workbench.

Topic : Diodes

Topic Objective:

At the end of this topic student would be able to:

- Understand diode operation and select diodes for various applications.
- Use the graphical load-line technique to analyze nonlinear circuits.
- Analyze and design simple voltage-regulator circuits.
- Use the ideal-diode model and piecewise-linear models to solve circuits.
- Understand various rectifier and wave-shaping circuits.
- Understand small-signal equivalent circuits.

Definition/Overview:

Diodes: In electronics, a diode is a two-terminal device (thermionic diodes may also have one or two ancillary terminals for a heater). Diodes have two active electrodes between which the signal of interest may flow, and most are used for their unidirectional electric current property. The varicap diode is used as an electrically adjustable capacitor.

Key Points:

1. Diode

In electronics, a diode is a two-terminal device (except that thermionic diodes may also have one or two ancillary terminals for a heater). Diodes have two active electrodes between which the signal of interest may flow, and most are used for their unidirectional current property. The varicap diode is used as an electrically adjustable capacitor.

The directionality of current flow most diodes exhibit is sometimes generically called the rectifying property. The most common function of a diode is to allow an electric current to pass in one direction (called the forward biased condition) and to block it in the opposite direction (the reverse biased condition). Thus, the diode can be thought of as an electronic version of a check valve. Real diodes do not display such a perfect on-off directionality but have a more complex non-linear electrical characteristic, which depends on the particular type of diode technology. Diodes also have many other functions in which they are not designed to operate in this on-off manner.

Early diodes included cats whisker crystals and vacuum tube devices (also called thermionic valves). Today the most common diodes are made from semiconductor materials such as silicon or germanium.

1.1. Thermionic and gaseous state diodes

Thermionic diodes are thermionic valve devices (also known as vacuum tubes), which are arrangements of electrodes surrounded by a vacuum within a glass envelope. Early examples were fairly similar in appearance to incandescent light bulbs.

In thermionic valve diodes, a current is passed through the heater filament. This indirectly heats the cathode, another filament treated with a mixture of barium and strontium oxides, which are oxides of alkaline earth metals; these substances are chosen because they have a small work function. (Some valves use direct heating, in which a tungsten filament acts as both cathode and emitter.) The heat causes thermionic emission of electrons into the vacuum. In forward operation, a surrounding metal electrode, called the anode, is positively charged, so that it electrostatically attracts the emitted electrons. However, electrons are not easily released from the unheated anode surface when the voltage polarity is reversed and hence any reverse flow is a very tiny current.

For much of the 20th century, thermionic valve diodes were used in analog signal applications, and as rectifiers in many power supplies. Today, valve diodes are only used in niche applications, such as rectifiers in guitar and hi-fi valve amplifiers, and specialized high-voltage equipment.

1.2. Semiconductor diodes

Most modern diodes are based on semiconductor p-n junctions. In a p-n diode, conventional current can flow from the p-type side (the anode) to the n-type side (the cathode), but cannot flow in the opposite direction. Another type of semiconductor diode, the Schottky diode, is formed from the contact between a metal and a semiconductor rather than by a p-n junction.

1.3. Zener Diodes

A Zener diode is a type of diode that permits current in the forward direction like a normal diode, but also in the reverse direction if the voltage is larger than the breakdown voltage known as "Zener knee voltage" or "Zener voltage". The device was named after Clarence Zener, who discovered this electrical property.

A conventional solid-state diode will not allow significant current if it is reverse-biased below its reverse breakdown voltage. When the reverse bias breakdown voltage is exceeded, a conventional diode is subject to high current due to avalanche breakdown. Unless this current is limited by external circuitry, the diode will be permanently damaged. In case of large forward bias (current in the direction of the arrow), the diode exhibits a voltage drop due to its junction built-in voltage and internal resistance. The amount of the voltage drop depends on the semiconductor material and the doping concentrations.

A Zener diode exhibits almost the same properties, except the device is specially designed so as to have a greatly reduced breakdown voltage, the so-called Zener voltage. A Zener diode contains a heavily doped p-n junction allowing electrons to tunnel from the valence band of the p-type material to the conduction band of the n-type material. In the atomic model, this tunneling corresponds to the ionization of covalent bonds. A reverse-biased Zener diode will exhibit a controlled breakdown and allow the current to keep the voltage across the Zener diode at the Zener voltage. For example, a diode with a Zener breakdown voltage of 3.2 V will exhibit a voltage drop of 3.2 V if reverse bias voltage applied across it is more than its Zener voltage. However, the current is not unlimited, so the Zener diode is typically used to generate a reference voltage for an amplifier stage, or as a voltage stabilizer for low-current applications.

2. Shockley diode equation

The Shockley ideal diode equation or the diode law (named after transistor co-inventor William Bradford Shockley, not to be confused with tetrode inventor Walter H. Schottky) is the IV characteristic of an ideal diode in either forward or reverse bias (or no bias). The equation is:

where

I is the diode current,

IS is a scale factor called the saturation current,

VD is the voltage across the diode,

VT is the thermal voltage,

and n is the emission coefficient, also known as the ideality factor. The emission coefficient n varies from about 1 to 2 depending on the fabrication process and semiconductor material and in many cases is assumed to be approximately equal to 1 (thus the notation n is omitted).

The thermal voltage VT is approximately 25.85 mV at 300 K, a temperature close to room temperature commonly used in device simulation software. At any temperature it is a known constant defined by:

Where

q is the magnitude of charge on an electron (the elementary charge),

k is Boltzmanns constant,

T is the absolute temperature of the p-n junction in kelvins

The Shockley ideal diode equation or the diode law is derived with the assumption that the only processes giving rise to current in the diode are drift (due to electrical field), diffusion, and thermal recombination-generation. It also assumes that the recombination-generation (R-G) current in the depletion region is insignificant. This means that the Shockley equation doesnt account for the processes involved in reverse breakdown and photon-assisted R-G. Additionally, it doesnt describe the leveling off of the IV curve at high forward bias due to internal resistance.

Under reverse bias voltages the exponential in the diode equation is negligible, and the current is a constant (negative) reverse current value of -IS. The reverse breakdown region is not modeled by the Shockley diode equation.

For even rather small forward bias voltages the exponential is very large because the thermal voltage is very small, so the subtracted 1 in the diode equation is negligible and the forward diode current is often approximated as

The use of the diode equation in circuit problems is illustrated in the article on diode modeling.

3. Assumed States for Analysis of Ideal-Diode Circuits

- Assume a state for each diode, either on (i.e., a short circuit) or off (i.e., an open circuit). For n diodes there are 2n possible combinations of diode states.

- Analyze the circuit to determine the current through the diodes assumed to be on and the voltage across the diodes assumed to be off.
- Check to see if the result is consistent with the assumed state for each diode. Current must flow in the forward direction for diodes assumed to be on. Furthermore, the voltage across the diodes assumed to be off must be positive at the cathode (i.e., reverse bias).
- If the results are consistent with the assumed states, the analysis is finished. Otherwise, return to step 1 and choose a different combination of diode states.

4. Piecewise linear model

In practice, the graphical method is complicated and long and is impractical for complex circuits. Another method of modeling a diode is called piecewise linear (PWL) modelling. In mathematics, this means taking a function and breaking it down into several linear segments. The example below shows how a curve can be approximated by three linear segments, forming a three-segment PWL model:

The same method is used to approximate the diode characteristic curve into linear segments. This enables us to substitute the real diode for an ideal diode, a voltage source and a resistor. The figure below shows a real diode I-V curve being approximated by a two segment piecewise linear model. Typically the sloped line segment would be chosen tangent to the diode curve at the Q-point. Then the slope of this line is given by the reciprocal of the small-signal resistance of the diode at the Q-point.

A piecewise linear approximation of the diode characteristic.

5. Peak Inverse Voltage

he peak inverse voltage is the specified maximum voltage that a diode rectifier will block.

5.1. In semiconductor diodes

As a general term applied to semiconductor diodes, peak reverse voltage or peak inverse voltage is the maximum voltage that a diode can withstand in the reverse direction without breaking down or avalanching. If this voltage is exceeded the diode may be destroyed. Diodes must have a peak inverse voltage rating that is higher than the maximum voltage that will be applied to them in a given application.

5.2. In rectifier applications

For rectifier applications, peak inverse voltage (PIV) or peak reverse voltage (PRV) is the maximum value of reverse voltage which occurs at the peak of the input cycle when the diode is reverse-biased. The portion of the sinusoidal waveform which does not repeat or duplicate itself is known as the cycle. The part of the cycle above the horizontal axis is called the positive half-cycle, or alternation; the part of the cycle below the horizontal axis is called the negative alternation. With reference to the amplitude of the cycle, the peak inverse voltage is specified as the maximum negative value of the sine-wave within a cycle's negative alternation.

6. Notation for Currents and Voltages in Electronic Circuits

- vD and iD represent the total instantaneous diode voltage and current. At times, we may wish to emphasize the time-varying nature of these quantities, and then we use vD(t) and iD(t)

- VDQ and IDQ represent the dc diode current and voltage at the quiescent point.
- vd and id represent the (small) ac signals. If we wish to emphasize their time varying nature, we use vd(t) and id(t).

Topic : Amplifiers: Specifications And External Characteristics

Topic Objective:

At the end of this topic student would be able to:

- Use various amplifier models to calculate amplifier performance for given sources and loads.
- Compute amplifier efficiency.
- Understand the importance of input and output impedances of amplifiers.
- Determine the best type of ideal amplifier for various applications.
- Specify the frequency-response requirements for various amplifier applications.
- Understand linear and nonlinear distortion in amplifiers.
- Specify the pulse-response parameters of amplifiers.
- Work with differential amplifiers and specify common-mode rejection requirements.
- Understand the various sources of dc offsets and design balancing circuits.

Definition/Overview:

Amplifier: Generally, an amplifier or simply amp, is any device that changes, usually increases, the amplitude of a signal. The "signal" is usually voltage or current. (So, an amplifier will generally increase the signal level)

Key Points:

1. Amplifier

Generally, an amplifier is any device that changes, usually increases, the amplitude of a signal. The "signal" is usually voltage or current. Ideally, an amplifier produces an output signal with identical waveshape as the input signal, but with a larger amplitude. In popular use, the term today usually refers to an electronic amplifier, often as in audio applications. The relationship of the input to the output of an amplifier usually expressed as a function of the input frequency is called the transfer function of the amplifier, and the magnitude of the transfer function is termed the gain. A related device that emphasizes conversion of signals of one type to another (for example, a light signal in photons to a DC signal in amperes) is a transducer, or a sensor. However, a transducer does not amplify power.

1.1. Inverting versus Noninverting Amplifiers

Inverting amplifiers have negative voltage gain, and the output waveform is an inverted version of the input waveform. Noninverting amplifiers have positive voltage gain.

2. Current Gain

In electronics, gain is a measure of the ability of a circuit (often an amplifier) to increase the power or amplitude of a signal. It is usually defined as the mean ratio of the signal output of a system to the signal input of the same system. It may also be defined as the decimal logarithm of the same ratio. Thus, the term gain on its own is ambiguous. For example, "a gain of five" may imply that either the voltage, current or the power is increased by a factor of five. Furthermore, the term gain is also applied in systems such as sensors where the input and output have different units; in such cases the gain units must be specified, as in "5 microvolts per photon" for a photosensor.

3. Power Gain

The power gain of an electrical network is the ratio of an output power to an input power. Unlike other signal gains, such as voltage and current gain, "power gain" may be ambiguous as the meaning of terms "input power" and "output power" is not always clear. Three important power gains are average power gain, transducer power gain and available power gain.

4. Cascaded Amplifiers

A cascade amplifier is any amplifier constructed from a series of amplifiers, where each amplifier sends its output to the input of the next amplifier in a daisy chain.

4.1. Simplified Models for Cascaded Amplifier Stages

First, determine the voltage gain of the first stage accounting for loading by the second stage. The overall voltage gain is the product of the gains of the separate stages. The input impedance is that of the first stage, and the output impedance is that of the last stage.

4.2. Importance Of Amplifier Impedances In Various Applications

Some applications call for amplifiers with high input (or output) impedance while others call for low input (or output) impedance. Other applications call for amplifiers that have specific input and/or output impedances.

3. Distortion

A distortion is the alteration of the original shape (or other characteristic) of an object, image, sound, waveform or other form of information or representation. Distortion is usually unwanted. In some fields, distortion is desirable, such as electric guitar (where distortion is often induced purposely with the amplifier or an electronic effect to achieve the electric guitar's desired, electrifying, aggressive sound). The slight distortion of analog tapes and vacuum tubes is considered pleasing in certain situations. The addition of noise or other extraneous signals (hum, interference) is not considered to be distortion, though the effects of distortion are sometimes considered noise.

3.1. Phase Distortion

In signal processing, phase distortion or phase-frequency distortion is distortion that occurs when (a) a filter's phase response is not linear over the frequency range of interest, that is, the phase shift introduced by a circuit or device is not directly proportional to frequency, or (b) the zero-frequency intercept of the phase-frequency characteristic is not 0 or an integral multiple of 2π radians.

3.2. Requirements for Distortionless Amplification

To avoid linear waveform distortion, an amplifier should have constant gain magnitude and a phase response that is linear versus frequency for the range of frequencies contained in the input signal.

3.3. Transfer Characteristic And Nonlinear Distortion

The transfer characteristic is a plot of instantaneous output amplitude versus instantaneous input amplitude. Curvature of the transfer characteristic results in nonlinear distortion.

3.4. Harmonic Distortion

Harmonic distortion is found in both the voltage and the current waveform. Most current distortion is generated by electronic loads, also called non-linear loads. These non-linear loads might be single phase loads such as point-of-sale terminals, or three-phase as in variable speed drives. As the current distortion is conducted through the normal system wiring, it creates voltage distortion according to Ohm's Law. While current distortion travels only along the power path of the non-linear load, voltage distortion affects all loads connected to that particular bus or phase. Current distortion affects the power system and distribution equipment. It may directly or indirectly cause the destruction of loads or loss of product. From the direct perspective, current distortion may cause transformers to overheat and fail even though they are not fully loaded. Conduc tors and conduit systems can also overheat leading to open circuits and downtime.

3.5. Total Harmonic Distortion

The total harmonic distortion, or THD, of a signal is a measurement of the harmonic distortion present and is defined as the ratio of the sum of the powers of all harmonic components to the power of the Fundamental frequency. Lesser THD, for example, allows the components in a loudspeaker, amplifier or microphone or other equipment to make a violin sound like a violin when played back, and not a cello or simply a distorted noise.

4. Differential Amplifiers

A differential amplifier is a type of electronic amplifier that multiplies the difference between two inputs by some constant factor (the differential gain). Many electronic devices use differential amplifiers internally. Given two inputs and , a practical differential amplifier gives an output V_{out}:

where A_d is the differential-mode gain and A_c is the common-mode gain.
The common-mode rejection ratio is usually defined as the ratio between differential-mode gain and common-mode gain:

In the above equation, as A_c approaches zero, CMRR approaches infinity. The higher the resistance of the current source R_e, the lower A_c is, and the better the CMRR. Thus, for a perfectly symmetrical differential amplifier with $A_c = 0$, the output voltage is given by:

Note that a differential amplifier is a more general form of amplifier than one with a single input; by grounding one input of a differential amplifier, a single-ended amplifier results. Some kinds of differential amplifier usually include several simpler differential amplifiers. For example, an instrumentation amplifier or a fully differential amplifier or a negative feedback amplifier or a instrument amplifier or a isolation amplifier often includes several op-amps; and those op-amps usually include a long-tailed pair.

A differential amplifier is the input stage of operational amplifiers, or op-amps, and emitter coupled logic gates.

Differential amplifiers are found in many systems that utilise negative feedback, where one input is used for the input signal, the other for the feedback signal. A common application is for the control of motors or servos, as well as for signal amplification applications. In discrete electronics, a common arrangement for implementing a differential amplifier is the long-tailed pair, which is also usually found as the differential element in most op-amp integrated circuits.

In Section 4 of this course you will cover these topics:

- Field-Effect Transistors
- Bipolar Junction Transistors
- Operational Amplifiers

Topic : Field-Effect Transistors

Topic Objective:

At the end of this topic student would be able to:

- Understand MOSFET operation.
- Use the load-line technique to analyze basic FET amplifiers.
- Analyze bias circuits.
- Use small-signal equivalent circuits to analyze FET amplifiers.
- Compute the performance parameters of several FET amplifier configurations.
- Select a FET amplifier configuration that is appropriate for a given application.
- Understand the basic operation of CMOS logic gates

Definition/Overview:

Field-Effect Transistors: The field-effect transistor (FET) is a type of transistor that relies on an electric field to control the shape and hence the conductivity of a 'channel' of one type of charge carrier in a semiconductor material. FETs are sometimes called unipolar transistors to contrast their single-carrier-type operation with the dual-carrier-type operation of bipolar (junction) transistors (BJT). The concept of the FET predates the BJT, though it was not physically implemented until after BJTs due to the limitations of semiconductor materials and relative ease of manufacturing BJTs compared to FETs at the time.

Key Points:

1. MOSFET

The metaloxidesemiconductor field-effect transistor (MOSFET, MOS-FET, or MOS FET) is a device used to amplify or switch electronic signals. It is by far the most common field-effect transistor in both digital and analog circuits. The MOSFET is composed of a channel of n-type or p-type semiconductor material (see article on semiconductor devices), and is accordingly called an NMOSFET or a PMOSFET (also commonly nMOSFET, pMOSFET). The 'metal' in the name is now often a misnomer because the previously metal gate material is now a layer of polysilicon (polycrystalline silicon; why polysilicon is used will be explained below). Previously aluminium was used as the gate material until the 1980s when polysilicon became dominant, owing to its capability to form self-aligned gates.

1.1. Operations

1.1.1. Metaloxidesemiconductor structure

A traditional metaloxidesemiconductor (MOS) structure is obtained by depositing a layer of silicon dioxide (SiO2) and a layer of metal (polycrystalline silicon is commonly used instead of metal) on top of a semiconductor die. As the silicon dioxide is a dielectric material its structure is equivalent to a planar capacitor, with one of the electrodes replaced by a semiconductor.

When a voltage is applied across a MOS structure, it modifies the distribution of charges in the semiconductor. If we consider a P-type semiconductor (with NA the density of acceptors, p the density of holes; p = NA in neutral bulk), a positive voltage, VGB, from gate to body (see figure) creates a depletion layer by forcing the positively charged holes away from the gate-insulator/semiconductor interface, leaving exposed a carrier-free region of immobile, negatively charged acceptor ions. See doping (semiconductor). If VGB is high enough, a high concentration of negative charge carriers forms in an inversion layer located in a thin layer next to the interface between the semiconductor and the insulator. (Unlike the MOSFET, discussed below, where the inversion layer electrons are supplied rapidly from the source/drain electrodes, in the MOS capacitor they are produced much more slowly by thermal generation through carrier generation and recombination centers in the depletion region.) Conventionally, the gate voltage at which the volume density of electrons in the inversion layer is the same as the volume density of holes in the body is called the threshold voltage. This structure with P-type body is the basis of the N-type MOSFET, which requires the addition of an N-type source and drain regions.

1.1.2. MOSFET structure and channel formation

A metaloxidesemiconductor field-effect transistor (MOSFET) is based on the modulation of charge concentration by a MOS capacitance between a body electrode and a gate electrode located above the body and insulated from all other device regions by a gate dielectric layer which in the case of a MOSFET is an oxide, such as silicon dioxide. If dielectrics other than an oxide such as silicon dioxide (often referred to as oxide) are employed the device may be referred to as a metalinsulatorsemiconductor FET (MISFET). The MOSFET includes two additional terminals (source and drain), each connected to individual highly doped regions that are separated by the body region. These regions can be either p or n type, but they must both be of the same type, and of opposite type to the body region. The highly doped source and drain regions typically are denoted by a '+' following the type of doping. The body is not highly doped, as denoted by the lack of a '+' sign.

If the MOSFET is an n-channel or nMOS FET, then the source and drain are 'n+' regions and the body is a 'p' region. As described above, with sufficient gate voltage, above a threshold voltage value, electrons from the source (and possibly also the drain) enter the inversion layer or n-channel at the interface between the p region and the oxide. This conducting channel extends between the source and the drain, and current is conducted through it when a voltage is applied between source and drain.

For gate voltages below the threshold value, the channel is lightly populated, and only a very small subthreshold leakage current can flow between the source and the drain. If the MOSFET is a p-channel or pMOS FET, then the source and drain are 'p+' regions and the body is a 'n' region. When a negative gate-source voltage (positive source-gate) is applied, it creates a p-channel at the surface of the n region, analogous to the n-channel case, but with opposite polarities of charges and voltages. When a voltage less negative than the threshold value (a negative voltage for p-Channel) is applied between gate and source, the channel disappears and only a very small subthreshold current can flow between the source and the drain.

The source is so named because it is the source of the charge carriers (electrons for n-channel, holes for p-channel) that flow through the channel; similarly, the drain is where the charge carriers leave the channel. The device may comprise a Silicon On Insulator (SOI) device in which a Buried OXide (BOX) is formed below a thin semiconductor layer. If the channel region between the gate dielectric and a Buried OXide (BOX) region is very thin, the very thin channel region is referred to as an Ultra Thin Channel (UTC) region with the source and drain regions formed on either side thereof in and/or above the thin semiconductor layer. Alternatively, the device may comprise a SEMiconductor On Insulator (SEMOI) device in which other semiconductors than silicon are employed. Many alternative semicondutor materials may be employed. When the source and drain regions are formed above the channel in whole or in part, they are referred to as Raised Source/Drain RSD) regions.

1.1.3. Modes of operation

The operation of a MOSFET can be separated into three different modes, depending on the voltages at the terminals. In the following discussion, a simplified algebraic model is used that is accurate only for old technology. Modern MOSFET characteristics require computer models that have rather more complex behavior.

- Cut-off or Sub-threshold or Weak Inversion Mode
- Triode Mode or Linear Region (also referred to as the Ohmic Mode
- Saturation Mode (also referred to as the Active Mode

2. NMOS and PMOS Transistors

nMOS logic uses n-type metal-oxide-semiconductor field effect transistors (MOSFETs) to implement logic gates and other digital circuits. nMOS transistors have three modes of operation: cut-off, triode, and saturation (sometimes called active).

The n-type MOSFETs are arranged in a so-called "pull-down network" (PDN) between the logic gate output and negative supply voltage, while a resistor is placed between the logic gate output and the positive supply voltage. The circuit is designed such that if the desired output is low, then the PDN will be active, creating a current path between the negative supply and the output.

As an example, here is a NOR gate in nMOS logic. If either input A or input B is high (logic 1, = True), the respective MOS transistor acts as a very low resistance between the output and the negative supply, forcing the output to be low (logic 0, = False). When both A and B are high, both transistors are conductive, creating an even lower resistance path to ground. The only case where the output is high is when both transistors are off, which occurs only when both A and B are low, thus satisfying the truth table of a NOR gate:

A	B	A NOR B
0	0	1
0	1	0
1	0	0
1	1	0

While nMOS logic is easy to design and manufacture (a MOSFET can be made to operate as a resistor, so the whole circuit can be made with nMOSFETs), it has several shortcomings as well. The worst problem is that a DC current flows through an nMOS logic gate when the PDN is active that is whenever the output is low. This leads to static power dissipation even when the circuit sits idle.

Also, nMOS circuits are slow to transition from low to high. When transitioning from high to low, the transistors provide low resistance, and the capacitative charge at the output drains away very quickly. But the resistance between the output and the positive supply rail is much greater, so the low to high transition takes longer. Using a resistor of lower value will speed up the process but also increases static power dissipation.

Additionally, the asymmetric input logic levels make nMOS circuits susceptible to noise.

These disadvantages are why nMOS logic was supplanted by CMOS logic both in low-power and in high-speed digital circuits, such as microprocessors, during the 1980s.

2.1. Load Line

A load line is used in graphic analysis of circuits; representing the constraint other parts of the circuit place on a non-linear device, like a diode or transistor. A load line represents the response of a resistor which shares a current with the device in question. Since both currents are the same, the operating point of the circuit will be at the intersection of the curve with the load line.

In the simple case of a diode shown, there is a single voltage across the diode and a single current through it. The load line represents the current in the resistor. When VD = VDD, there will be no voltage across the resistor, so the current will be 0. If VD = 0, the current will be at its maximum.

In a BJT circuit, the BJT has a different current-voltage(IC-VCE) characteristic depending on the Base current. Placing a series of these curves on the graph shows how the base current will affect the operating point of the circuit.

It should be noted that the load line is used for dc analysis, and has no bearing on small-signal analysis once an operating point is identified.

3. Common-Source Amplifiers

In electronics, a common-source amplifier is one of three basic single-stage field-effect transistor (FET) amplifier topologies, typically used as a voltage or transconductance amplifier. The easiest way to tell if a FET is common source, common drain, or common gate is to examine where the signal enters, and leaves. The remaining terminal is what is known as "common". In this example, the signal enters the gate, and exits the drain. The only terminal remaining is the source. This is a common-source FET circuit. The analogous bipolar junction transistor circuit is the common-emitter amplifier.

The common-source (CS) amplifier may be viewed as a transconductance amplifier or as a voltage amplifier. (See classification of amplifiers). As a transconductance amplifier, the input voltage is seen as modulating the current going to the load. As a voltage amplifier, input voltage modulates the amount of current flowing through the FET, changing the voltage across the output resistance according to Ohm's law. However, the FET device's output resistance typically is not high enough for a reasonable transconductance amplifier (ideally infinite), nor low enough for a decent voltage amplifier (ideally zero). Another major drawback is the amplifier's limited high-frequency response. Therefore, in practice the output often is routed through either a voltage follower (common-drain or CD stage), or a current follower (common-gate or CG stage), to obtain more favorable output and frequency characteristics. The CSCG combination is called a cascode amplifier.

3.1. The Small-Signal Equivalent Circuit

In small-signal midband analysis of FET amplifiers, the coupling capacitors, bypass capacitors, and dc voltage sources are replaced by short circuits. The FET is replaced with its small-signal equivalent circuit. Then, we write circuit equations and derive useful expressions for gains, input impedance, and output impedance.

4. CMOS

Complementary metaloxidesemiconductor (CMOS), is a major class of integrated circuits. CMOS technology is used in microprocessors, microcontrollers, static RAM, and other digital logic circuits. CMOS technology is also used for a wide variety of analog circuits such as image sensors, data converters, and highly integrated transceivers for many types of communication. Frank Wanlass successfully patented CMOS in 1967 (US Patent 3,356,858).

CMOS was also sometimes referred to as complementary-symmetry metaloxidesemiconductor (or COS-MOS). The words "complementary-symmetry" refer to the fact that the typical digital design style with CMOS uses complementary and symmetrical pairs of p-type and n-type metal oxide semiconductor field effect transistors (MOSFETs) for logic functions.

Two important characteristics of CMOS devices are high noise immunity and low static power consumption. Significant power is only drawn when the transistors in the CMOS device are switching between on and off states. Consequently, CMOS devices do not produce as much waste heat as other forms of logic, for example transistor-transistor logic (TTL) or NMOS logic, which uses all n-channel devices without p-channel devices. CMOS also allows a high density of logic functions on a chip.

The phrase "metaloxidesemiconductor" is a reference to the physical structure of certain field-effect transistors, having a metal gate electrode placed on top of an oxide insulator, which in turn is on top of a semiconductor material. Instead of metal (usually aluminum in the very old days), current gate electrodes (including those up to the 65 nanometer technology node) are almost always made from a different material, polysilicon, but the terms MOS and CMOS nevertheless continue to be used for the modern descendants of the original process.

Topic : Bipolar Junction Transistors

Topic Objective:

At the end of this topic student would be able to:

- Understand bipolar junction transistor operation in amplifier circuits.
- Use the load-line technique to analyze simple amplifiers and understand the causes of nonlinear distortion.
- Use large-signal equivalent circuits to analyze BJT circuits.
- Analyze bias circuits.
- Use small-signal equivalent circuits to analyze BJT amplifiers.
- Compute performance of several important amplifier configurations.
- Select an amplifier configuration appropriate for a given application.

Definition/Overview:

Bipolar Junction Transistors: A bipolar (junction) transistor (BJT) is a type of transistor. It is a three-terminal device constructed of doped semiconductor material and may be used in amplifying or switching applications. Bipolar transistors are so named because their operation involves both electrons and holes, as opposed to unipolar transistors, such as field-effect transistors, in which only one carrier type is involved in charge flow.

Key Points:

1.Bipolar Junction Transistor

1.1. Introduction

An NPN transistor can be considered as two diodes with a shared anode region. In typical operation, the emitterbase junction is forward biased and the basecollector junction is reverse biased. In an NPN transistor, for example, when a positive voltage is applied to the baseemitter junction, the equilibrium between thermally generated carriers and the repelling electric field of the depletion region becomes unbalanced, allowing thermally excited electrons to inject into the base region. These electrons wander (or "diffuse") through the base from the region of high concentration near the emitter towards the region of low concentration near the collector. The electrons in the base are called minority carriers because the base is doped p-type which would make holes the majority carrier in the base.

The base region of the transistor must be made thin, so that carriers can diffuse across it in much less time than the semiconductor's minority carrier lifetime, to minimize the percentage of carriers that recombine before reaching the collectorbase junction. To ensure this, the thickness of the base is much less than the diffusion length of the electrons. The collectorbase junction is reverse-biased, so little electron injection occurs from the collector to the base, but electrons that diffuse through the base towards the collector are swept into the collector by the electric field in the depletion region of the collectorbase junction.

1.2. Voltage, Current, and Charge Control

The collectoremitter current can be viewed as being controlled by the baseemitter current (current control), or by the baseemitter voltage (voltage control). These views are related by the currentvoltage relation of the baseemitter junction, which is just the usual exponential currentvoltage curve of a p-n junction (diode).

The physical explanation for collector current is the amount of minority-carrier charge in the base region. Detailed models of transistor action, such as the GummelPoon model, account for the distribution of this charge explicitly to explain transistor behavior more exactly. The charge-control view easily handles photo-transistors, where minority carriers in the base region are created by the absorption of photons, and handles the dynamics of turn-off, or recovery time, which depends on charge in the base region recombining. However, since base charge is not a signal that is visible at the terminals, the current- and voltage-control views are usually used in circuit design and analysis.

In analog circuit design, the current-control view is sometimes used since it is approximately linear. That is, the collector current is approximately βF times the base current. Some basic circuits can be designed by assuming that the emitterbase voltage is approximately constant, and that collector current is beta times the base current. However, to accurately and reliably design production bjt circuits, the voltage-control (for example, EbersMoll) model is required The voltage-control model requires an exponential function to be taken into account, but when it is linearized such that the transistor can be modelled as a transconductance, as in the EbersMoll model, design for circuits such as differential amplifiers again becomes a mostly linear problem, so the voltage-control view is often preferred. For translinear circuits, in which the exponential IV curve is key to the operation, the transistors are usually modelled as voltage controlled with transconductance proportional to collector current. In general, transistor level circuit design is performed using SPICE or a comparable analogue circuit simulator, so model complexity is usually not of much concern to the designer.

1.3. Transistor 'Alpha' and 'Beta'

The proportion of electrons able to cross the base and reach the collector is a measure of the BJT efficiency. The heavy doping of the emitter region and light doping of the base region cause many more electrons to be injected from the emitter into the base than holes to be injected from the base into the emitter. The common-emitter current gain is represented by βF or hfe. It is approximately the ratio of the DC collector current to the DC base current in forward-active region, and is typically greater than 100. Another important parameter is the common-base current gain, αF. The common-base current gain is approximately the gain of current from emitter to collector in the forward-active region. This ratio usually has a value close to unity; between 0.98 and 0.998. Alpha and beta are more precisely related by the following identities (NPN transistor):

1.4. Structure

A BJT consists of three differently doped semiconductor regions, the emitter region, the base region and the collector region. These regions are, respectively, p type, n type and p type in a PNP, and n type, p type and n type in a NPN transistor. Each semiconductor region is connected to a terminal, appropriately labeled: emitter (E), base (B) and collector (C).

The base is physically located between the emitter and the collector and is made from lightly doped, high resistivity material. The collector surrounds the emitter region, making it almost impossible for the electrons injected into the base region to escape being collected, thus making the resulting value of α very close to unity, and so, giving the transistor a large β. A cross section view of a BJT indicates that the collectorbase junction has a much larger area than the emitterbase junction.

The bipolar junction transistor, unlike other transistors, is usually not a symmetrical device. This means that interchanging the collector and the emitter makes the transistor leave the forward active mode and start to operate in reverse mode. Because the transistor's internal structure is usually optimized to forward-mode operation, interchanging the collector and the emitter makes the values of α and β in reverse operation much smaller than those found in forward operation; often the α of the reverse mode is lower than 0.5. The lack of symmetry is primarily due to the doping ratios of the emitter and the collector. The emitter is heavily doped, while the collector is lightly doped, allowing a large reverse bias voltage to be applied before the collectorbase junction breaks down. The collectorbase junction is reverse biased in normal operation. The reason the emitter is heavily doped is to increase the emitter injection efficiency: the ratio of carriers injected by the emitter to those injected by the base. For high current gain, most of the carriers injected into the emitterbase junction must come from the emitter.

The low-performance "lateral" bipolar transistors sometimes used in CMOS processes are sometimes designed symmetrically, that is, with no difference between forward and backward operation.

Small changes in the voltage applied across the baseemitter terminals causes the current that flows between the emitter and the collector to change significantly. This effect can be used to amplify the input voltage or current. BJTs can be thought of as voltage-controlled current sources, but are more simply characterized as current-controlled current sources, or current amplifiers, due to the low impedance at the base.

Early transistors were made from germanium but most modern BJTs are made from silicon. A significant minority are also now made from gallium arsenide, especially for very high speed applications

1.5. NPN

NPN is one of the two types of bipolar transistors, in which the letters "N" and "P" refer to the majority charge carriers inside the different regions of the transistor. Most bipolar transistors used today are NPN, because electron mobility is higher than hole mobility in semiconductors, allowing greater currents and faster operation.

NPN transistors consist of a layer of P-doped semiconductor (the "base") between two N-doped layers. A small current entering the base in common-emitter mode is amplified in the collector output. In other terms, an NPN transistor is "on" when its base is pulled high relative to the emitter.

The arrow in the NPN transistor symbol is on the emitter leg and points in the direction of the conventional current flow when the device is in forward active mode.

One mnemonic device for identifying the symbol for the NPN transistor is "not pointing in".

1.5. PNP

The other type of BJT is the PNP with the letters "P" and "N" referring to the majority charge carriers inside the different regions of the transistor.

PNP transistors consist of a layer of N-doped semiconductor between two layers of P-doped material. A small current leaving the base in common-emitter mode is amplified in the collector output. In other terms, a PNP transistor is "on" when its base is pulled low relative to the emitter.

The arrow in the PNP transistor symbol is on the emitter leg and points in the direction of the conventional current flow when the device is in forward active mode.

One mnemonic device for identifying the symbol for the PNP transistor is "points in proudly".

1.6. Heterojunction Bipolar Transistor

The heterojunction bipolar transistor (HBT) is an improvement of the BJT that can handle signals of very high frequencies up to several hundred GHz. It is common nowadays in ultrafast circuits, mostly RF systems. Heterojunction transistors have different semiconductors for the elements of the transistor. Usually the emitter is composed of a larger bandgap material than the base. The figure shows that this difference in bandgap allows the barrier for holes to inject backward into the base, denoted in figure as $\Delta\varphi p$, to be made large, while the barrier for electrons to inject into the base $\Delta\varphi n$ is made low. This barrier arrangement helps reduce minority carrier injection from the base when the emitter-base junction is under forward bias, and thus reduces base current and increases emitter injection efficiency.

The improved injection of carriers into the base allows the base to have a higher doping level, resulting in lower resistance to access the base electrode. In the more traditional BJT, also referred to as homojunction BJT, the efficiency of carrier injection from the emitter to the base is primarily determined by the doping ratio between the emitter and base, which means the base must be lightly doped to obtain high injection efficiency, making its resistance relatively high. In addition, higher doping in the base can improve figures of merit like the Early voltage by lessening base narrowing.

The grading of composition in the base, for example, by progressively increasing the amount of germanium in a SiGe transistor, causes a gradient in bandgap in the neutral base, denoted in the figure by $\Delta\varphi G$, providing a "built-in" field that assists electron transport across the base. That drift component of transport aids the normal diffusive transport, increasing the frequency response of the transistor by shortening the transit time across the base.

Two commonly used HBTs are silicongermanium and aluminum gallium arsenide, though a wide variety of semiconductors may be used for the HBT structure. HBT structures are usually grown by epitaxy techniques like MOCVD and MBE.

2. Transistors in Circuits

The diagram opposite is a schematic representation of an npn transistor connected to two voltage sources. To make the transistor conduct appreciable current (on the order of 1 mA) from C to E, VBE must be above a minimum value sometimes referred to as the cut-in voltage. The cut-in voltage is usually about 600 mV for silicon BJTs, but can be different depending on the current level selected for the application and the type of transistor. This applied voltage causes the lower p-n junction to 'turn-on' allowing a flow of electrons from the emitter into the base. Because of the electric field existing between base and collector (caused by VCE), the majority of these electrons cross the upper p-n junction into the collector to form the collector current, IC. The remainder of the electrons recombine with holes, the majority carriers in the base, making a current through the base connection to form the base current, IB . As shown in the diagram, the emitter current, IE, is the total transistor current which is the sum of the other terminal currents. That is:

In the diagram, the arrows representing current point in the direction of the electric or conventional currentthe flow of electrons is in the opposite direction of the arrows since electrons carry negative electric charge. The ratio of the collector current to the base current is called the DC current gain. This gain is usually quite large and is often 100 or more.

It should also be noted that the emitter current is related to VBE exponentially. At room temperature, increasing VBE by about 60 mV increases the emitter current by a factor of 10. The base current is approximately proportional to the emitter current, so it varies the same way.

3. Regions of Operation

Bipolar transistors have five distinct regions of operation, defined mostly by applied bias:

Forward-active (or simply, active): The emitter-base junction is forward biased and the base-collector junction is reverse biased. Most bipolar transistors are designed to afford the greatest common-emitter current gain, βf, in forward-active mode. If this is the case, the collector-emitter current is approximately proportional to the base current, but many times larger, for small base current variations.

Reverse-active (or inverse-active or inverted): By reversing the biasing conditions of the forward-active region, a bipolar transistor goes into reverse-active mode. In this mode, the emitter and collector regions switch roles. Since most BJTs are designed to maximize current gain in forward-active mode, the βf in inverted mode is several (2 - 3 for the ordinary germanium transistor) times smaller. This transistor mode is seldom used, usually being considered only for failsafe conditions and some types of bipolar logic. The reverse bias breakdown voltage to the base may be an order of magnitude lower in this region.

Saturation: With both junctions forward-biased, a BJT is in saturation mode and facilitates high current conduction from the emitter to the collector. This mode corresponds to a logical "on", or a closed switch.

Cutoff: In cutoff, biasing conditions opposite of saturation (both junctions reverse biased) are present. There is very little current flow, which corresponds to a logical "off", or an open switch.

3.1. Avalanche Breakdown Region

While these regions are well defined for sufficiently large applied voltage, they overlap somewhat for small (less than a few hundred millivolts) biases. For example, in the typical grounded-emitter configuration of an NPN BJT used as a pulldown switch in digital logic, the "off" state never involves a reverse-biased junction because the base voltage never goes below ground; nevertheless the forward bias is close enough to zero that essentially no current flows, so this end of the forward active region can be regarded as the cutoff region.

4. Large-Signal Models

The DC emitter and collector currents in active mode are well modeled by an approximation to the EbersMoll model:

The base internal current is mainly by diffusion and

Where

VT is the thermal voltage kT / q (approximately 26 mV at 300 K ≈ room temperature).

IE is the emitter current

IC is the collector current

αT is the common base forward short circuit current gain (0.98 to 0.998)

IES is the reverse saturation current of the baseemitter diode (on the order of 10−15 to 10−12 amperes)

VBE is the baseemitter voltage

Dn is the diffusion constant for electrons in the p-type base

W is the base width

The collector current is slightly less than the emitter current, since the value of αT is very close to 1.0. In the BJT a small amount of baseemitter current causes a larger amount of collectoremitter current. The ratio of the allowed collectoremitter current to the baseemitter current is called current gain, β or hFE. A β value of 100 is typical for small bipolar transistors. In a typical configuration, a very small signal current flows through the baseemitter junction to control the emittercollector current. β is related to α through the following relations:

Emitter Efficiency :; that is, the ratio of current injected into the base to the current in the emitter; the two differ due to backward injection from the base into the emitter and to recombination.

The unapproximated Ebers-Moll equations used to describe the three currents in any operating region are given below. These equations are based on the transport model for a bipolar junction transistor.

where

iC is the collector current

iB is the base current

iE is the emitter current

βF is the forward common emitter current gain (20 to 500)

βR is the reverse common emitter current gain (0 to 20)

IS is the reverse saturation current (on the order of 10^{-15} to 10^{-12} amperes)

VT is the thermal voltage (approximately 26 mV at 300 K ≈ room temperature).

VBE is the baseemitter voltage

VBC is the basecollector voltage

Topic : Operational Amplifiers

Topic Objective:

At the end of this topic student would be able to:

- List the characteristics of ideal op amps.
- Identify negative feedback in op-amp circuits.
- Use the summing-point constraint to analyze ideal op-amp circuits that have negative feedback.

- Select op-amp circuit configurations suitable for various applications.
- Use op amps to design useful circuits.
- Identify practical op-amp limitations and recognize potential inaccuracies in instrumentation applications.
- Work with instrumentation amplifiers.
- Applyintegrators, differentiators, and activefilters.

Definition/Overview:

Operational Amplifiers: An operational amplifier, often called an op-amp , is a DC-coupled high-gain electronic voltage amplifier with differential inputs and, usually, a single output. Typically the output of the op-amp is controlled either by negative feedback, which largely determines the magnitude of its output voltage gain, or by positive feedback, which facilitates regenerative gain and oscillation. High input impedance at the input terminals and low output impedance are important typical characteristics.

Key Points:

1. Ideal Operational Amplifier

Shown on the right is an example of an ideal operational amplifier. The main part in an amplifier is the dependent voltage source that increases in relation to the voltage drop across Rin, thus amplifying the voltage difference between V+ and V-. Many uses have been found for operational amplifiers and an ideal op-amp seeks to characterize the physical phenomena that make op-amps useful.

Vs+ and Vs- are not connected to the circuit within the op-amp because they power the dependent voltage source's circuit (not shown). These are notable, however, because they determine the maximum voltage the dependent voltage source can output.

For any input voltages the ideal op-amp has

- infinite open-loop gain,
- infinite bandwidth,
- infinite input impedances (resulting in zero input currents),
- zero offset voltage,
- infinite slew rate,
- zero output impedance, and
- zero noise.

Because of these properties, an op-amp can be modeled as a nullor.

1.1. Characteristics of Ideal Op Amps

- o Infinite gain for the differential input signal
- o Zero gain for the common-mode input signal
- o Infinite input impedance
- o Zero output impedance
- o Infinite bandwidth

1.2. Summing-Point Constraint

Operational amplifiers are almost always used with negative feedback, in which part of the output signal is returned to the input in opposition to the source signal. In a negative feedback system, the ideal op-amp output voltage attains the value needed to force

the differential input voltage and input current to zero. We call this fact the summing-point constraint.

1.3. Ideal Op-Amp Circuits

Ideal Op-Amp Circuits are analyzed by the following steps:

- Verify that negative feedback is present.
- Assume that the differential input voltage and the input current of the op amp are forced to zero. (This is the summing-point constraint.)
- Apply standard circuit-analysis principles, such as Kirchhoffs laws and Ohms law, to solve for the quantities of interest.

1.4. Circuit notation

The circuit symbol for an op-amp is shown in Figure 1

where:

V+: non-inverting input

V−: inverting input

Vout: output

VS+: positive power supply

VS−: negative power supply

The power supply pins (VS+ and VS−) can be labeled in different ways. Despite different labeling, the function remains the same. Often these pins are left out of the diagram for clarity, and the power configuration is described or assumed from the circuit.

1.5. Operation of ideal op-amps

The amplifier's differential inputs consist of an inverting input and a non-inverting input, and ideally the op-amp amplifies only the difference in voltage between the two. This is called the "differential input voltage". In its most common use, the op-amp's output voltage is controlled by feeding a fraction of the output signal back to the inverting input. This is known as negative feedback. If that fraction is zero, i.e., there is no negative feedback, the amplifier is said to be running "open loop" and its output is the differential input voltage multiplied by the total gain of the amplifier, as shown by the following equation:

where V+ is the voltage at the non-inverting terminal, V− is the voltage at the inverting terminal and G is the total open-loop gain of the amplifier.

Because the magnitude of the open-loop gain is typically very large and not well controlled by the manufacturing process, op-amps are not usually used without negative feedback. Unless the differential input voltage is extremely small, open-loop operation results in op-amp saturation. An example of how the output voltage is calculated when negative feedback exists is shown below in Basic non-inverting amplifier circuit.

Another typical configuration of op-amps is the positive feedback, which takes a fraction of the output signal back to the non-inverting input. An important application of it is the comparator with hysteresis.

For any input voltages the ideal op-amp has

infinite open-loop gain,

infinite bandwidth,

infinite input impedances (resulting in zero input currents),

zero offset voltage,

infinite slew rate,

zero output impedance, and

zero noise.

The inputs of an ideal op-amp under negative feedback can be modeled using a nullator, the output with a norator and the combination (complete ideal op-amp) by a nullor.

1.6. Limitations of real op-amps

Real op-amps can only approach this ideal: in addition to the practical limitations on slew rate, bandwidth, offset and so forth mentioned above, real op-amp parameters are subject to drift over time and with changes in temperature, input conditions, etc. Modern integrated FET or MOSFET op-amps approximate more closely the ideal op-amp than bipolar ICs where large signals must be handled at room temperature over a limited bandwidth; input impedance, in particular, is much higher, although the bipolar op-amps usually exhibit superior (i.e., lower) input offset drift and noise characteristics.

Where the limitations of real devices can be ignored, an op-amp can be viewed as a black box with gain; circuit function and parameters are determined by feedback, usually negative. IC op-amps as implemented in practice are moderately complex integrated circuits.

2. Inverting Amplifier

Inverts and amplifies a voltage (multiplies by a negative constant)

- $Z_{in} = R_{in}$ (because V_{-} is a virtual ground)
- A third resistor, of value , added between the non-inverting input and ground, while not necessary, minimizes errors due to input bias currents.

3. Non-Inverting amplifier

Amplifies a voltage (multiplies by a constant greater than 1)

- (realistically, at least the input impedance of the opamp itself, 1 MΩ to 10 TΩ. In many cases, the input impedance is significantly higher as a consequence of the feedback network)

- A third resistor, of value , added between the V_{in} source and the non-inverting input, while not necessary, minimizes errors due to input bias currents.

4. Voltage Followers

Used as a buffer amplifier, to eliminate loading effects or to interface impedances (connecting a device with a high source impedance to a device with a low input impedance). Due to the strong feedback, this circuit tends to get unstable when driving a high capacity load. This can be avoided by connecting the load through a resistor.

- (realistically, the differential input impedance of the op-amp itself, 1 MΩ to 1 TΩ)

5. Nonlinear Limitations

The output voltage of a real op amp is limited to the range between certain limits that depend on the internal design of the op amp. When the output voltage tries to exceed these limits, clipping occurs.

5.1. Slew-Rate Limitation

Another nonlinear limitation of actual op amps is that the magnitude of the rate of change of the output voltage is limited.

5.2. Full-Power Bandwidth

The full-power bandwidth of an op amp is the range of frequencies for which the op amp can produce an undistorted sinusoidal output with peak amplitude equal to the guaranteed maximum output voltage.

6. DC Imperfections

The three dc imperfections (bias current, offset current, and offset voltage) can be modeled by placing dc sources at the input of the op amp as shown in Figure 2. The effect of bias current, offset current, and offset voltage on inverting or noninverting amplifiers is to add a (usually undesirable) dc

voltage to the intended output signal.

7. Differential and Instrumentation Amplifiers

7.1. Differential Amplifiers

The circuit shown is used for finding the difference of two voltages each multiplied by some constant (determined by the resistors).
The name "differential amplifier" should not be confused with the "differentiator", also shown on this page.

Differential Z_{in} (between the two input pins) = $R_1 + R_2$

7.2. Instrumentation Amplifiers

Combines very high input impedance, high common-mode rejection, low DC offset, and other properties used in making very accurate, low-noise measurements. Is made by adding a non-inverting buffer to each input of the differential amplifier to increase the input impedance.

7.3. Integrators and Differentiators

Integrators produce output voltages that are proportional to the running time integral of the input voltages. In a running time integral, the upper limit of integration is t .

8. Active Filters

Filters can be very useful in separating desired signals from noise. Ideally, an active filter circuit should:

- Contain few components
- Have a transfer function that is insensitive to component tolerances
- Place modest demands on the op amps gainbandwidth product, output impedance, slew rate, and other specifications
- Be easily adjusted
- Require a small spread of component values
- Allow a wide range of useful transfer functions to be realized

In Section 5 of this course you will cover these topics:

- Magnetic Circuits And Transformers
- Dc Machines
- Ac Machines

Topic : Magnetic Circuits And Transformers

Topic Objective:

At the end of this topic student would be able to:

- Understand magnetic fields and their interactions with moving charges.
- Use the right-hand rule to determine the direction of the magnetic field around a current-carrying wire or coil.
- Calculate forces on moving charges and currentcarrying wires due to magnetic fields.
- Calculate the voltage induced in a coil by a changingmagnetic flux or in a conductor cutting through a magnetic field.
- Use Lenzs law to determine the polarities of induced voltages.
- Apply magnetic-circuit concepts to determine the magnetic fields in practical devices.
- Determine the inductance and mutual inductance of coils, given their physical parameters.
- Understand hysteresis, saturation, core loss, and eddy currents in cores composed of magnetic materials such as iron.
- Understand ideal transformers and solve circuits that include transformers.
- Use the equivalent circuits of real transformers to determine their regulations and power efficiencies.

Definition/Overview:

Magnetic Circuits: A magnetic circuit is a closed path containing a magnetic flux. It generally contains magnetic elements such as permanent magnets, ferromagnetic materials, and electromagnets, but may also contain air gaps and other materials.

Transformers: A transformer is a device that transfers electrical energy from one circuit to another through inductively coupled electrical conductors. A changing current in the first circuit (the primary) creates a changing magnetic field; in turn, this magnetic field induces a changing voltage in the second circuit (the secondary); this is called mutual induction. By adding a load to the secondary circuit, one can make current flow in the transformer, thus transferring energy from one circuit to the other.

Key Points:

1. Magnetic Fields

In physics, a magnetic field is a vector field that permeates space and which can exert a magnetic force on moving electric charges and on magnetic dipoles (such as permanent magnets). When placed in a magnetic field, magnetic dipoles tend to align their axes to be parallel with the magnetic field, as can be seen when iron filings are in the presence of a magnet (see picture at right). In addition, a changing magnetic field can induce an electric field. Magnetic fields surround and are created by electric currents, magnetic dipoles, and changing electric fields. Magnetic fields also have their own energy, with an energy density proportional to the square of the field intensity.

There are some notable specific instances of the magnetic field. For the physics of magnetic materials, see magnetism and magnet, and more specifically ferromagnetism, paramagnetism, and diamagnetism. For constant magnetic fields, such as are generated by stationary dipoles and steady currents, see magnetostatics. For magnetic fields created by changing electric fields, see electromagnetism.

The electric field and the magnetic field are tightly interlinked, in two senses. First, changes in either of these fields can cause ("induce") changes in the other, according to Maxwell's equations. Second, according to Einstein's theory of special relativity, a magnetic force in one inertial frame of reference may be an electric force in another, or vice-versa (see relativistic electromagnetism for examples). Together, these two fields make up the electromagnetic field, which is best known for underlying light and other electromagnetic waves.

1.1. Right-Hand Rule

In mathematics and physics, the right-hand rule is a common mnemonic for understanding notation conventions for vectors in 3 dimensions. When choosing three vectors that must be at right angles to each other, there are two distinct solutions, so when expressing this idea in mathematics, one must remove the ambiguity of which solution is meant. There are variations on the mnemonic depending on context, but all variations are related to the one idea of choosing a convention.

1.2. Lenzs Law

Lenzs law states that the polarity of the induced voltage is such that the voltage

would produce a current (through an external resistance) that opposes the original change in flux linkages.

2. Magnetic Circuits

In many engineering applications, we need to compute the magnetic fields for structures that lack sufficient symmetry for straight-forward application of Ampres law. Then, we use an approximate method known as magnetic-circuit analysis.

2.1. Advantage of the Magnetic-Circuit Approach

The advantage of the magnetic-circuit approach is that it can be applied to unsymmetrical magnetic cores with multiple coils.

3. Transformers

A key application of transformers is to increase voltage before transmitting electrical energy over long distances through wires. Wires have resistance and so dissipate electrical energy at a rate proportional to the square of the current through the wire. By transforming electrical power to a high-voltage (and therefore low-current) form for transmission and back again afterwards, transformers enable economic transmission of power over long distances. Consequently, transformers have shaped the electricity supply industry, permitting generation to be located remotely from points of demand. All but a tiny fraction of the world's electrical power has passed through a series of transformers by the time it reaches the consumer.Transformers are used extensively in electronic products to step down the supply voltage to a level suitable for the low voltage circuits they contain. The transformer also electrically isolates the end user from contact with the supply voltage.

Signal and audio transformers are used to couple stages of amplifiers and to match devices such as microphones and record player cartridges to the input impedance of amplifiers. Audio transformers allowed telephone circuits to carry on a two-way conversation over a single pair of wires. Transformers are also used when it is necessary to couple a differential-mode signal to a ground-referenced signal, and for isolation between external cables and internal circuits.

3.1. Basic principles

The transformer is based on two principles: firstly, that an electric current can produce a magnetic field (electromagnetism) and secondly that a changing magnetic field within a coil of wire induces a voltage across the ends of the coil (electromagnetic induction). By changing the current in the primary coil, it changes the strength of its magnetic field; since the changing magnetic field extends into the secondary coil, a voltage is induced across the secondary. A simplified transformer design is shown to the left. A current passing through the primary coil creates a magnetic field. The primary and secondary coils are wrapped around a core of very high magnetic permeability, such as iron; this ensures that most of the magnetic field lines produced by the primary current are within the iron and pass through the secondary coil as well as the primary coil.

3.1.1. Induction law

The voltage induced across the secondary coil may be calculated from Faraday's law of induction, which states that:

where V_S is the instantaneous voltage, N_S is the number of turns in the secondary coil and Φ equals the magnetic flux through one turn of the coil. If the turns of the coil are oriented perpendicular to the magnetic field lines, the flux is the product of the magnetic field strength B and the area A through which it cuts. The area is constant, being equal to the cross-sectional area of the transformer core, whereas the magnetic field varies with time according to the excitation of the primary. Since the same magnetic flux passes through both the primary and secondary coils in an ideal transformer,[1] the instantaneous voltage across the primary winding equals

Taking the ratio of the two equations for V_S and V_P gives the basic equation[5] for stepping up or stepping down the voltage

Topic : Dc Machines

Topic Objective:

At the end of this topic student would be able to:

- Select the proper motor type for various applications.
- State how torque varies with speed for various motors.
- Use the equivalent circuit for dc motors to compute electrical and mechanical quantities.
- Use motor nameplate data.
- Understand the operation and characteristics of shunt-connected dc motors, series-connected dc motors, and universal motors.

Definition/Overview:

DC Machine: The effect of the commutator is to produce a fixed spatial distribution of current directions in the armature conductors (shown as blue & green circles) independent of shaft rotation. The field created by these currents (armature reaction) is vertically directed along the quadrature axis. The field established by the excitation of the stator poles is directed along the horizontal direct axis.

Key Points:

1. DC Machines

The field windings provide the excitation necessary to set up the magnetic fields in the machine. There are various types of field windings that can be used in the generator or motor circuit. In addition to the following field winding types, permanent magnet fields are used on some smaller DC products.

1.1. Separately Excited

Separately Excited Winding When the field is connected to an external power source, it is a separately excited field...

1.2. Shunt Winding

Shunt wound motors, with the armature shunted across the field, offer relatively flat speed-torque characteristics. Combined with inherently controlled no-load speed, this provides good speed regulation over wide load ranges. While the starting torque is comparatively lower than the other DC winding types, shunt wound motors offer simplified control for reversing service.

Shunt windings usually consists of a large number of turns of small size wire. The torque/ current curve is non-linear above full load. Shunt wound motors often have a rising speed characteristic with increased load.

1.3. Series Winding

Series wound motors have the armature connected in series with the field. While it offers very high starting torque and good torque output per ampere, the series motor has poor speed regulation. Speed of DC series motors is generally limited to 5000 rpm and below. These are generally used on crane and hoist applications.

A series winding usually consists of a small number of turns of large size wire. With this winding, the motor can produce high starting and overload torque. This design is not used for applications with light loads or no load conditions. Series motors should be avoided in applications where they are likely to lose there load because of their tendency to "run away" under no-load conditions.

1.3. Compound Winding

Compound wound (stabilized shunt) motors utilize a field winding in series with the armature in addition to the shunt field to obtain a compromise in performance between a series and shunt type motor. This type offers a combination of good starting torque and speed stability. Standard compounding is about 12%. Heavier compounding of up to 40 to 50% can be supplied for special high starting torque applications, such as hoists and cranes.

2. Commutation

The maximum voltage from an armature winding can be obtained when the brushes are in contact with those conductors, which are midway between the poles. This will result in the greatest possible number of conductors cutting the magnetic lines in one direction between a positive and negative brush. This brush position is known as the no load neutral position of the brushes.

The current in a given armature coil reverses in direction as the coil sides move from one pole to another of opposite polarity, whereas the function of the commutator is to keep the current unidirectional. This reversal of current is known as commutation. The commutator acts as a switch to keep the current flowing in one direction. However, the fast rate of change in direction of the current in any given coil induces an appreciable voltage in that coil which tends to keep the current flowing in the original direction. Therefore, the current reversal is delayed causing an accelerated rate of change near the end of the commutation period. This results in an arc if the reversal is not completed before the brush breaks contact with the coil involved. Any arcing is detrimental to the operation of the machine and must be counteracted.

3. Speed Control OF DC Motors

- Vary the voltage supplied to the armature circuit while holding the field constant.
- Vary the field current while holding the armature supply voltage constant.
- Insert resistance in series with the armature circuit.

Topic : Ac Machines

Topic Objective:

At the end of this topic student would be able to:

- Select the proper ac motor type for various applications.
- State how torque varies with speed for various ac motors.
- Compute electrical and mechanical quantities for ac motors.
- Use motor nameplate data.

- Understand the operation and characteristics of three-phase induction motors, three-phase synchronous machines, and various types of single-phase ac motors, stepper motors, and brushless dc motors.

Definition/Overview:

AC Motors: An AC motor is an electric motor that is driven by an alternating current. It consists of two basic parts, an outside stationary stator having coils supplied with AC current to produce a rotating magnetic field, and an inside rotor attached to the output shaft that is given a torque by the rotating field. There are two types of AC motors, depending on the type of rotor used. The first is the synchronous motor, which rotates exactly at the supply frequency or a submultiple of the supply frequency. The magnetic field on the rotor is either generated by current delivered through slip rings or by a permanent magnet. The second type is the induction motor, which turns slightly slower than the supply frequency. The magnetic field on the rotor of this motor is created by an induced current.

Key Points:

1. Induction Motor

An induction motor (IM) is a type of asynchronous AC motor where power is supplied to the rotating device by means of electromagnetic induction. An electric motor converts electrical power to mechanical power in its rotor (rotating part). There are several ways to supply power to the rotor. In a DC motor this power is supplied to the armature directly from a DC source. But in an AC motor this power is induced in the rotating device. An induction motor can be called a rotating transformer because the stator (stationary part) is essentially the primary side of the transformer and the rotor (rotating part) is the secondary side. Induction motors are widely used, especially polyphase induction motors, which are frequently used in industrial drives.

Induction motors are now the preferred choice for industrial motors due to their rugged construction, lack of brushes and thanks to modern power electronics the ability to control the speed of the motor.

2. Principle of operation and comparison to synchronous motors

The basic difference between an induction motor and a synchronous AC motor is that in the latter a current is supplied onto the rotor. This then creates a magnetic field which, through magnetic interaction, links to the rotating magnetic field in the stator which in turn causes the rotor to turn. It is called synchronous because at steady state the speed of the rotor is the same as the speed of the rotating magnetic field in the stator.

By way of contrast, the induction motor does not have any direct supply onto the rotor; instead, a secondary current is induced in the rotor. To achieve this, stator windings are arranged around the rotor so that when energised with a polyphase supply they create a rotating magnetic field pattern which sweeps past the rotor. This changing magnetic field pattern can induce currents in the rotor conductors. These currents interact with the rotating magnetic field created by the stator and the rotor will turn.

However, for these currents to be induced, the speed of the physical rotor and the speed of the rotating magnetic field in the stator must be different, or else the magnetic field will not be moving relative to the rotor conductors and no currents will be induced. If by some chance this happens, the rotor typically slows slightly until a current is re-induced and then the rotor continues as before. This difference between the speed of the rotor and speed of the rotating magnetic field in the stator is called slip. It is unitless and is the ratio between the relative speed of the magnetic field as seen by the rotor to the speed of the rotating field. Due to this an induction motor is sometimes referred to as an asynchronous machine.

3. Construction

The stator consists of wound 'poles' that carry the supply current that induces a magnetic field in the conductor. The number of 'poles' can vary between motor types but the poles are always in pairs (i.e. 2,4,6 etc). There are two types of rotor:

- Squirrel-cage rotor
- Slip ring rotor

The most common rotor is a squirrel-cage rotor. It is made up of bars of either solid copper (most common) or aluminum that span the length of the rotor, and are connected through a ring at each end. The rotor bars in squirrel-cage induction motors are not straight, but have some skew to reduce noise and harmonics.

The motor's phase type is one of two types:

- Single-phase induction motor
- 3-phase induction motor

4. Speed control

The rotational speed of the rotor is controlled by the number of pole pairs (number of windings in the stator) and by the frequency of the supply voltage. Before the development of cheap power electronics, it was difficult to vary the frequency to the motor and therefore the uses for the induction motor were limited.

There are various techniques to produce a desired speed. The most commonly used technique is PWM (Pulse Width Modulation), in which a DC signal is switched on and off very rapidly, producing a sequence of electrical pulses to the inductor windings. The duty cycle of the pulses, also known as the mark-space ratio, determines the average power input to the motor. For example, a 100 V DC signal that is cut into on- and off- pulses of equal width, has an average voltage of 50 V. If the on- pulses are a third of the duration of the off pulses, the average would be 25 V. The frequency of the pulses determines the motor speed.

The general term for a power electronic device that controls the speed as well as other parameters is called an 'inverter'. A typical unit will take the mains AC supply, rectify and smooth it into a "link" DC voltage, and, by using the method described above, converts it into the desired AC waveform.

Because the induction motor has no brushes and is easy to control, many older DC motors are being replaced with induction motors and accompanying inverters in industrial applications.

5. Starting of induction motor

In a three phase induction motor, the induced emf in the rotor circuit depends on the slip of the induction motor and the magnitude of the rotor current depends upon this induced emf (electromotive force). When the motor is started, the slip is equal to 1 as the rotor speed is zero, so the induced emf in the rotor is large. As a result, a very high current flows through the rotor. This is similar to a transformer with the secondary coil short circuited, which causes the primary coil to draw a high current from the mains. Similarly, when an induction motor starts, a very high current is drawn by the stator, on the order of 5 to 9 times the full load current. This high current can damage the motor windings and because it causes heavy line voltage drop, other appliances connected to the same line may be affected by the voltage fluctuation. To avoid such effects, the starting current should be limited. A soft start starter is a device which limits the starting current by providing reduced voltage to the motor. Once the rotor speed increases, the full rated voltage is given to it.

6. Single Phase Electric Power

In electrical engineering, single-phase electric power refers to the distribution of alternating current electric power using a system in which all the voltages of the supply vary in unison. Single-phase distribution is used when loads are mostly lighting and heating, with few large electric motors. A single-phase supply connected to an alternating current electric motor does not produce a revolving magnetic field; single-phase motors need additional circuits for starting, and such motors are uncommon above 10 or 20 kW in rating.

In contrast, in a three-phase system, the currents in each conductor reach their peak instantaneous values sequentially, not simultaneously; in each cycle of the power frequency, first one, then the second, then the third current reaches its maximum value. The waveforms of the three supply conductors are offset from one another in time (delayed in phase) by one-third of their period.

Standard frequencies of single-phase power systems are either 50 or 60 Hz. Special single-phase traction power networks may operate at 16.67 Hz or other frequencies to power electric railways.

7. Phase versus Line Quantities

The voltage Vs across each winding and current Is through each winding shown in Figure 17.13 are called the phase voltage and phase current, respectively. The windings of an induction motor may be connected in either a delta or a wye.