Essential Theory for the Electronics Hobbyist

G.T. RUBAROE, T.Eng.(CEI), Assoc.I.E.R.E.

ESSENTIAL THEORY FOR THE ELECTRONICS HOBBYIST

ESSENTIAL THEORY FOR THE ELECTRONICS HOBBYIST

by

G.T. RUBAROE, T.Eng.(CEI), Assoc.IERE

BERNARDS (Publishers) LTD
The Grampians
Shepherds Bush Road
London W6 7NF
England

I.S.B.N. 0 900162 69 4

First Published September 1977

Printed and Manufactured in Great Britain by
C. Nicholls & Co. Ltd.

CONTENTS

Page

Chapter 1

INTRODUCTION

Electronics is very popular with the general public as a leisure activity. Much of the enthusiasm has been generated due to two reasons. Firstly, improving technology and mass production methods have made it possible to offer to the hobbyist a wide variety of semiconductor and other components at very competitive prices. Secondly, excellent books and monthly publications are readily available, and these provide detailed information concerning projects. These also provide valuable technical guidance so important to the hobbyist, for the successful completion of any project.

In any hobby activity a background knowledge of the subject can increase considerably the employment and satisfaction one derives from it. This viewpoint applies without any reservations whatsoever, to electronics. Electronics as a subject can be treated at many levels. Most of the books available at the present time treat the subject with an academic slant, and for use as instruction material by students preparing for examinations. Treatment at this level is of limited value to the hobbyist whose primary interest is to build a given circuit and make it work. It is in this context that a background knowledge tailored to meet his specific requirements can come in useful. In this book the author has brought together the relevant material and presented it in a readable manner with minimum recourse to mathematics. A smattering of arithmetic and basic algebra is all that is needed to follow the subject material.

Electronics is intimately associated with atomic theory. The operation and behaviour of every electronic component depends in one way or another on atomic phenomena. Therefore a feeling for what goes on in the atomic world is a prerequisite to a better understanding of these components. The second chapter is devoted to this topic. The reader is introduced to atomic structure and this is followed by an exposition of the electron.

An understanding of electrical conduction in solids is of significant importance in modern electronics, which is largely dominated by solid-state devices. The theory relating to this topic is examined in Chapter 3. We draw heavily on the material covered in Chapter 2 to describe the conduction mechanisms in conductors and semiconductors, and to distinguish between conductors, semiconductors, and insulators. The current flow in semiconductor devices takes place due to the presence of two types of charge carriers, namely the electron and the hole. These concepts are briefly discussed in view of their relevance to device operation.

In Chapter 4 we undertake a discussion on electrical parameters used to describe the behaviour of circuits. The concepts connected with these parameters are developed from basic reasoning.

All components used in electronic circuits can be broadly grouped into linear and non-linear components. In Chapter 5 and Chapter 6 we cover the properties unique to these components. The behaviour of dc and ac circuits are investigated in Chapter 7. The section on ac circuits includes a brief discussion on signal representation. Chapter 8 is devoted to diode and transistor circuits. Biasing methods are discussed and attention is drawn to popular circuit applications. In recent years we have witnessed the elevation of integrated circuits to a position of great importance. These remarkable devices are used widely both in the analogue and digital areas. In Chapter 9 we examine the salient features of these devices and take a look at some of the popular applications.

A book on essential theory will be incomplete without at least a passing reference to measurements. The main function of any electronic measurement is to identify some aspect of circuit behaviour. For the results to be of any value they need to be accurate and meaningful. Unless one keeps a cautious watch over the entire measurement situation, various sources of error can influence the results. The final chapter draws the reader's attention to common pitfalls and briefly discusses a few typical measurements.

Many formulae having a practical bearing are presented in this book. Purpose-designed examples are employed to illustrate their application. These examples are presented at the end of each chapter.

The electronic calculator can be very helpful in removing much of the drudgery of routine computations. A simple single-memory, four-function machine will be adequate for most day-to-day requirements. The time involved in carrying out certain calculations can be considerably reduced by handling formulae in a manner which allows a chain calculation to be performed. This approach often reduces the number of key operations and also eliminates the need for elaborate memory facilities, or pencil and paper to record intermediate steps.

Chapter 2

THE ATOM AND ITS CONSTITUENTS

In our daily affairs we have become accustomed to seeing things in terms of size, shape, colour and so on. The faculty of vision involves the eye. Now, there is a point to remember here; that is there is a limit to the smallest size of object that the eye can resolve or see clearly. As we move on to still smaller dimensions, our faculty of vision simply disappears. Of course, we have combated this problem to a certain degree by designing both optical and electron microscopes. When we go still deeper into this microscopic world, in fact far beyond the reaches of the most powerful microscopes we have, we are in the so-called ultra microscopic world of the atom. Atomic Physics is a very fascinating area of science, and has the potential power to either make or mar human destiny. The concepts that have been developed over the last 100 years or so to explain and understand atomic phenomena have served to explain numerous other phenomena which occur at macroscopic levels. The term macroscopic is used to describe large things, easily visible to the eye.

Electronics derives its name from one of the constituents of the atom, namely the electron. It is due to this reason that any discussion on electronics is incomplete without reference to the ultra microscopic world of the atom. The basic substances in our environment are known as elements. All other substances are made up from combinations of two or more of these elements. The atom is the smallest particle that can exist in nature without losing its identity as an element. Elements are classified according to their atomic structure. Each element has a structure which is unique to itself. The atom may be regarded as the fundamental building brick of nature.

Let us now examine the structure of the atom, more closely. The particles that constitute the atom are the electron, proton, and neutron. There are two properties relating to these particles, of special interest to us. These concern their mass and electrical charge. Mass refers to the quantity of matter. The unit of mass in SI (Systeme Internationale) units is the kilogram. The electron has a mass which is extremely small. Stretching our imagination to any extent will not help us to visualize its minuteness. The mass of the proton is about 1840 times that of the electron, and still happens to be extremely small. The neutron has a mass equal to that of the proton. The property of charge refers to the quantity of electricity. The unit of charge is the coulomb. The electron carries a negative charge while the proton carries a positive charge of equal value. The neutron on the other hand carries no charge.

The electrons, protons and neutrons are arranged in a very interesting way to form the atomic structure. The protons and neutrons make up what is known as the nucleus. The electrons are visualized as being in continual motion around the nucleus. The solar system serves as an useful analogy in this context. The sun corresponds to the nucleus and the planets correspond to the orbiting electrons. Scaling the solar system down to ultra microscopic dimensions gives us a model which describes the situation in a very simplified manner.

The simplest atom in nature is that of hydrogen. Its nucleus consists of one proton. The requirement for equal protons and electrons is met by the presence of a single orbiting electron. Figure 2.1 shows the structure of the hydrogen atom.

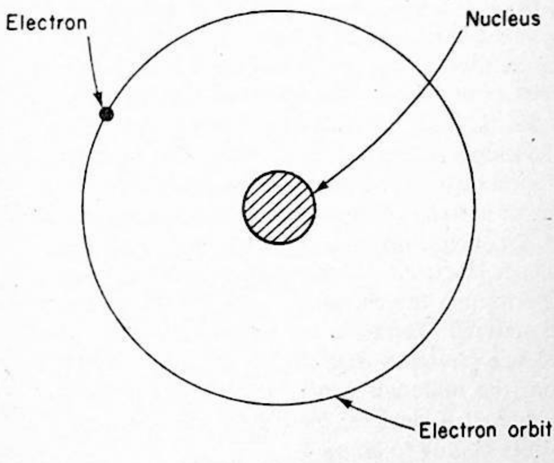


Fig. 2.1 Structure of hydrogen atom.

The elements may be arranged according to the number of electrons their respective atoms contain, in an ascending order. In this way an element could be identified in terms of its atomic number. The hydrogen atom has one orbiting electron and hence its atomic number is 1. Next to hydrogen comes helium which has two orbiting electrons. Its atomic number is therefore 2. As the atomic number increases the structure becomes very complex.

The planetory model of the atom was first proposed by Rutherford. It was later perfected by Bohr, making use of Planck's quantum theory. In Bohr's model the electrons which rotate around the nucleus are permitted to do so only at given distances from the nucleus. The exact radial positions depend upon the energy possessed by the electrons. Energy is defined as the capacity to do work. It is measured in joules. When dealing with electrons and other charged particles it is convenient to use another unit. This is the electron-volt. The electron-volt is related to the joule by a simple relationship and can be easily converted when necessary. As the number of electrons surrounding the nucleus increase, they are grouped into what are known as shells and sub-shells. The shell structure is illustrated in Figure 2.2. In this figure only the shells are shown. The point to note is that the shells in turn consists of sub-shells. The shells are identified by the capital letters K, L, M, N, O, P, and Q. The sub-shells are identified by the simple letters s, p, d, f, and g. The number of electrons that each shell or sub-shell can accommodate is fixed. In atoms of certain elements the outer shell is not filled to capacity. Electrons in an unfilled outer shell are known as valence electrons. These electrons play an important role in determining the chemical and electrical properties of a given material. Elements are grouped according to the number of valence electrons their atoms contain. Germanium and silicon, the materials widely used in the manufacture of semiconductor devices, have four valence electrons and therefore belong to group 4.

So far we have considered an isolated atom. It is interesting to examine the state of affairs when the atom is just one of many millions which participate in building up a sample of a given element. Naturally there will have to be some force which act to keep the atoms together. Scientists have identified three types of bonding mechanisms. These are called metallic, ionic and covalent bonding. In electronics our main

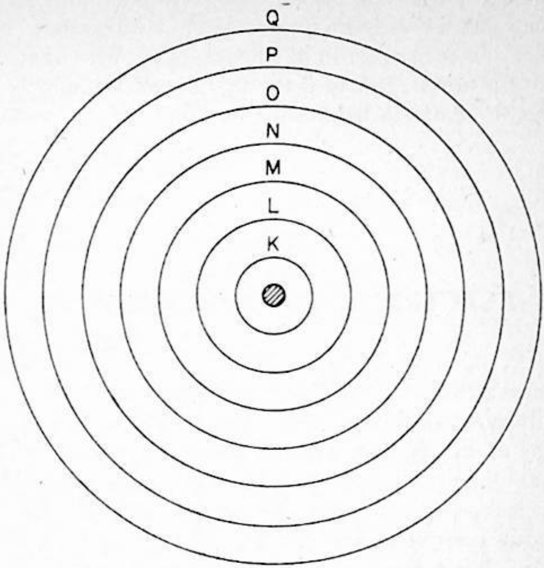


Fig. 2.2 Arrangement of shells in an atom.

interest is in metallic and covalent bonding. Metallic bonding as the name suggests applies to the bonding of metals, as for example copper. Covalent bonding applies to semiconductor materials such as germanium and silicon. The four valence electrons in the outermost shell function to keep the atoms together. Both, with metallic and covalent bonding the resulting structure exhibits the properties of a crystal. A crystal structure is characterized by the manner in which the atoms arrange to produce a regular geometric pattern.

From experience we are aware that the electrical properties of materials can differ in many respects. Some are good conductors of electricity while others are not. Then we have the semiconductor materials which exhibit conducting properties mid-way between the two extreme cases of conductors and insulators. It is well known that the conducting properties of materials are affected by changes in temperature. Materials differ from each other in this respect too. What is so elegant about the atomic theory is the way in which it clearly explains to us why materials behave the way they do.

Chapter 3

ELECTRICAL CONDUCTION IN SOLIDS

Whenever there is charge movement, we have a situation which signifies electrical conduction. We have already seen that two of the atom's constituents are charged particles. These are namely the proton and electron. The charge carried by these particles are equal in value and opposite in sign. The proton carries a positive charge and the electron carries a negative charge. When the number of electrons exactly equal the number of protons the atom is said to be electrically neutral. If for some reason the atom loses an electron it is called a positive ion. On the other hand if it gains an electron we have what is known as a negative ion. Since the atom is no longer electrically neutral under these conditions it could be looked upon as a charged particle. Therefore any movement of ions also amount to electrical conduction. The mass of an ion is many times that of an electron. Due to this reason it will be relatively immobile.

A discussion of electrical conduction in solids can be undertaken by considering independently the conduction mechanisms in conductors and semiconductors.

Any material which allows charge movement to take place without hindrance is known as an electrical conductor. Copper, iron and aluminium are materials which come under this category. Atomic theory helps us to understand why these and many others behave as conductors. Good conductors of electricity are normally metals. The atoms of a metal are held together by the mechanism of metallic bonding. To really see what takes place let us take as an example the copper atom. A pictorial representation of the copper atom is shown in Figure 3.1. It is evident from the figure that this

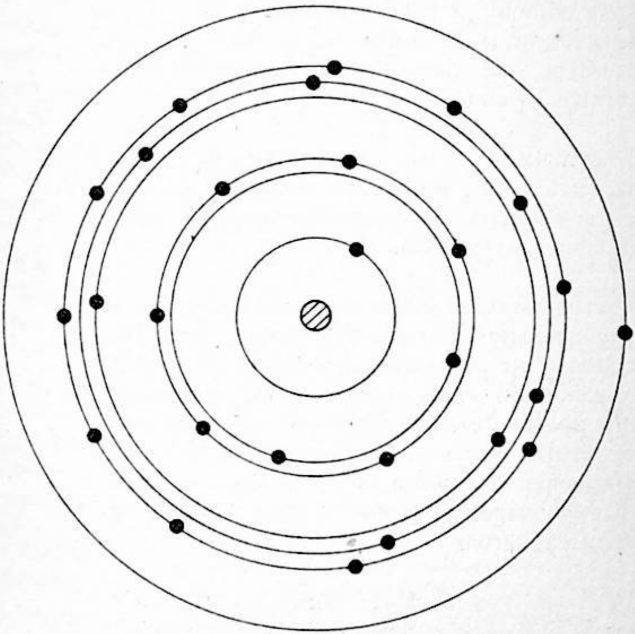


Fig. 3.1 Structure of copper atom.

atom has 29 electrons orbiting around its nucleus. The outer N shell contains just one electron. In terms of atomic dimensions this electron is placed quite far away from the nucleus. Hence the nucleus has very little influence over it. Now suppose we have a situation where the copper atom is surrounded by millions of other copper atoms. This and all the other atoms will be held together due to a continual exchange between atoms of the single electron in the outer N shell of each atom. This electron never spends much time with any particular atom and is always on the move. Binding of atoms in this manner is known as metallic bonding. It is important to realize that the copper atom remains electrically neutral at all times. Although the outer electron keeps changing places all the time, at no time is any given atom allowed to be without the full complement of 29 electrons required to compensate the positive charge associated with the nucleus. It will be remembered that this positive charge is due to the protons which reside in the nucleus.

What makes a material a good conductor is the presence of the wandering electrons referred to above. An application of a directed electric force immediately causes these electrons to move in a preferred direction.

Let us now examine how electrical conduction takes place in semiconductor materials. The most important semiconductor materials are germanium and silicon. These materials are elements and belong to group 4. We have already seen that it is possible to classify elements according to the number of valence electrons, into groups. When we say that germanium and silicon belong to group 4, we simply mean that the respective atoms of these elements have four valence electrons.

Only one type of charge carrier, namely the electron is responsible for electrical conduction in conductor materials. With semiconductor materials, the conduction processes involve two types of charge carrier, namely the electron and hole. The hole is the name given to the vacancy created by an electron breaking away from the parent atom. In a way the hole may be looked upon as a ficticious entity, whose presence is necessary to explain the conduction processes associated with semiconductor materials.

Before considering the conduction processes, it is instructive to examine how the atoms are held together in germanium and silicon. In this respect what is true for germanium is equally true for silicon. Therefore we are justified in confining our treatment to one type. Silicon has been chosen due to its wider application as a semiconductor material. The structure of a silicon atom is shown in Figure 3.2. It has 14

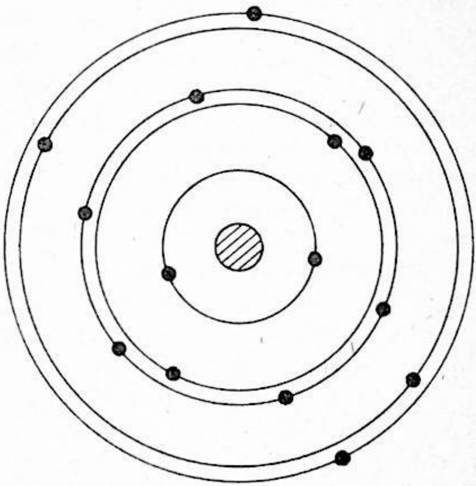


Fig. 3.2 Structure of silicon atom.

orbiting electrons, of which 4 are in the unfilled M shell. In a sample of silicon there are millions of atoms packed together to form a crystal structure, and in close proximity to each other. In this structure, the valence electrons of a given atom while being attracted by its own nucleus comes also under the influence of neighbouring nuclei. This gives rise to a sharing process and forms the basis on which the atoms are bound to each other. When valence electrons take part in binding atoms together in the manner described, the process is known as covalent bonding.

In pure or intrinsic silicon, the regular geometrical array characteristic to a crystalline substance is perfect only at zero degrees kelvin. This is a temperature reference familiar to the physicist. In degrees centigrade zero degrees kelvin is equal to −273. Absolute zero is often used in place of zero degrees kelvin and carries the same meaning. At absolute zero temperature, the atoms of the crystal lattice is stationary. At temperatures above absolute zero, the thermal energy of the environment causes the atoms to vibrate about their mean positions. This in turn causes a disruption in the covalent bonding which serves to hold the atoms together. The rupture of covalent bonds in this manner produces electrons and holes in equal number. These have to be equal since by definition a hole is a vacancy in the crystal lattice caused by an electron leaving a parent atom. The electron-hole pairs behave as charge carriers and bestow upon the material, semiconducting properties. The production of electron-hole pairs increases with temperature and this explains why semiconductor devices are temperature sensitive.

Silicon in its pure form with an equal number of electrons and holes is not useful in the manufacture of semiconductor devices. It is necessary somehow to get the material into a state where one type of charge carrier, that is the electron or hole predominate. This is exactly what happens when extremely small quantities of group 3 or group 5 materials are added to pure silicon under carefully controlled conditions. The materials added, are called impurities, and the process is known as doping. Adding of group 3 materials produces what is known as p-type silicon, while adding of group 5 materials produce what is known as n-type silicon. In p-type silicon the predominant carrier is the hole, whereas in n-type silicon it is the electron. The presence of these charge carriers allow electrical conduction to take place in these materials. Depending upon the type of material, conduction within it will be due either to the electron or hole.

We have seen that electrical conduction in solids is made possible by the presence of charge carriers. By the same reasoning any material depleted of charge carriers will not allow electrical conduction to take place. Materials in this category are called insulators. Conducting, semiconducting, and insulating materials are all equally important in electronics.

Chapter 4

ELECTRICAL PARAMETERS

Any quantity which can be employed to characterize the structure and operation of an electrical circuit is called an electrical parameter. This is a very broad definition and embraces both constant and variable quantities. The constant parameters of interest to us are resistance, inductance and capacitance. In this context we must realize that these quantities are liable to vary in conditions where significant changes in the environment occur. The variable parameters of interest to us are current, voltage, power, impedance, and admittance. The two variable parameters, current and voltage may be regarded as the basic ones from which the others are expressed.

Constant Parameters

Resistance

From our discussion of atomic structure and electrical conduction in solids, we know that a material consists of atoms packed closely together. We also know that the electrons in the outer shells can behave as charge carriers. The resistance parameter refers to the opposition experienced by these charge carriers when they endeavour to move within the material. They are unable to move very far before colliding with an atom. This gives rise to a loss of energy which manifests itself as heat in the material. For a given charge movement the number of collisions increases with resistance. Therefore the heat lost in the material would also increase. Resistance to movement of charge is a fundamental physical property of all materials and is a constant under given environmental conditions. In attempting to quantify any quantity we require units of measurement. The unit of measurement of resistance is the ohm.

Inductance

Wherever there is charge movement there is also an associated magnetic field. So far as the rate of charge movement remains constant the strength of the magnetic field too remains constant. However the moment a change in the rate of charge movement occurs, the magnetic field has to immediately readjust itself to meet the new situation. This readjustment functions to oppose the change in the rate of charge movement. This opposition to the rate of charge movement has to be accounted for in some way, and the inductance parameter does just that. The unit of inductance is the henry (H). The henry is a very large unit and in practice we often prefer to work in terms of millihenrys (mH) or microhenrys (μH). A millihenry is 1 thousandth part of a henry. A microhenry is 1 millionth part of a henry.

Capacitance

Wherever there is charge present there is an associated electric field. Any other charge placed in this electric field experiences a force. This phenomenon has to be accounted for when characterizing a circuit. The capacitance parameter does exactly this. The unit of capacitance is the farad (F). The farad like the henry is a very large unit. In practice we often use smaller units, namely the microfarad (μF), picofarad (pF), and nanofarad (nF). The microfarad is 1 millionth part of a farad. The picofarad is 1 millionth part of a microfarad, or 1 millionth, millionth part of a farad. The nanofarad is 1 thousandth part of a microfarad, or 1 millionth, thousandth part of a farad.

Variable Parameters

The variable parameters we shall discuss are current, voltage, power, impedance, and admittance. However before doing so we need to take a brief look at the methods employed for representing variable quantities. In any discussion involving variable quantities it is useful to distinguish between a dependent and an independent variable. An independent variable is one which varies quite independently and is in no way dependent on another. Time is a good example of an independent variable. A dependent variable on the other hand is one that varies entirely under the command of another. It enjoys no independence whatsoever.

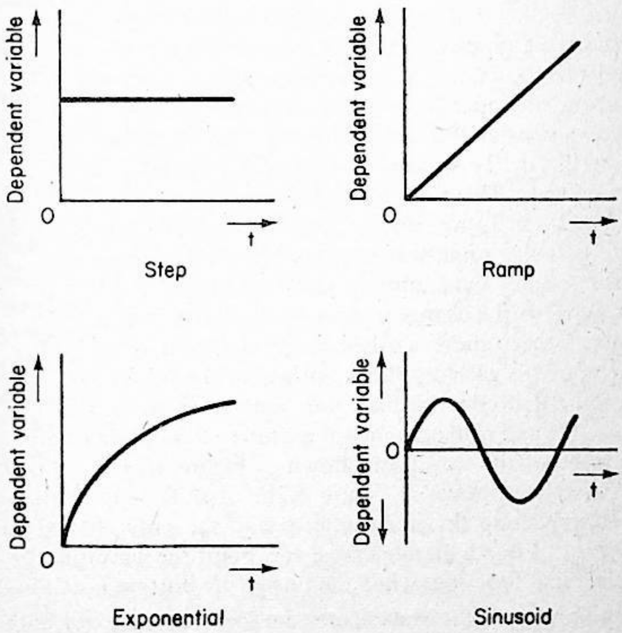


Fig. 4.1 Commonly occurring waveforms.

The step function, ramp, exponential, and sinusoid are waveforms frequently encountered when investigating electronic circuits, and are associated with the variable parameters which describe the circuit behaviour. These waveforms are shown in Figure 4.1. In diagrams of this type showing the relationship between an independent and dependent variable it is usual to show the independent variable along the horizontal axis and the dependent variable along the vertical axis. In all four diagrams the independent variable is shown to be time as is often the case in practical situations. A step function is one which changes from zero to a given value in zero time. We have a practical example of a step function type of excitation whenever we close a switch to connect a battery to a circuit. A ramp function is one which varies linearly with time. The exponential function is based on the mathematical constant e. e has a value 2.718 and is the base of natural logarithms. The sinusoid is a function which varies in a manner whereby the dependent variable bears a trigonometric relationship to the independent variable. The importance of this waveform in its role as a variable function cannot be over-emphasized. Certain properties of the sinusoid need mentioning while we are on this subject. These are highlighted in Figure 4.2. The displacement along the vertical axis, is called the amplitude. Note that the amplitude changes are cyclic. Starting from zero it reaches a maximum positive value, and then continues to change with a change in slope and reaches zero again after a time T. The time T is called the period of the sinusoid. The frequency f is given by the reciprocal of the period. In referring to amplitude, we may use terms such as maximum amplitude and instantaneous amplitude. If we were to plot the slope of the waveform shown in Figure 4.2a we obtain the waveform shown in Figure 4.2b. A shift in the entire waveform along the horizontal axis is the only change obsefved. From a mathematical viewpoint the waveform on top is a sine function, while the one at the bottom is a cosine function. Since the waveshapes are identical they are both grouped under the one name of sinusoid. Taking the slope of any waveform yields its differential. We say that the function has been differentiated. It will thus be seen that the derivative

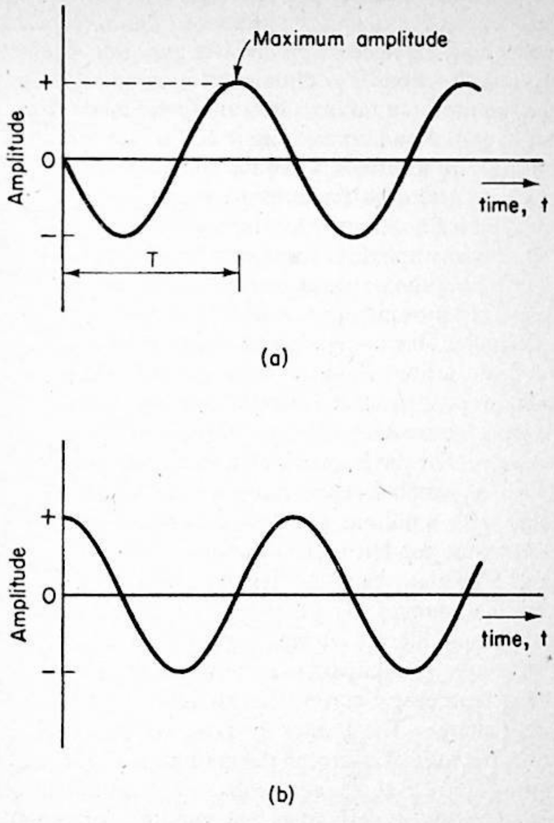


Fig. 4.2 The important features of the sinusoid.

of a sine function is a cosine function, having as we have already mentioned exactly the same waveshape. In later work we will see that this property has a major bearing for the popularity of the sinusoid as a test waveform, in certain applications.

Current
Movement of charge in a directed fashion constitute a current. In an electrically conducting material the predominant charge carrier is the electron, and the presence of an electrical current wthin the material necessarily involves a motion of electrons in a preferred direction. For directed motion an electromotive force is required. In taking a length of wire made from a conducting material and connecting it across the terminals of a battery, we are in effect satisfying this requirement. Since the current flowing in the external circuit serves no useful purpose and is limited only by the internal resistance of the battery, this example is of value only in a theoretical sense. It demonstrates the need for an electromotive force to establish a directed movement of charge. The unit of current is the ampere (A). This being a fairly large unit we often use smaller units derived from it. These are the milliampere (mA), the microampere (μA), and the nanoampere (nA). These units are derived by dividing an ampere by 10^3, 10^6, and 10^9, respectively. For the benefit of those readers not familiar with this form of number representation 10^3 simply means a thousand, 10^6, a million, and 10^9, a thousand million. The numeric on the top left of 10 is known as the exponent and indicates how many times 10 has to be taken as a factor to give the final number. The basic unit of charge is the charge on an electron. Since a current is in effect a directed movement of charge, the ampere as an unit has to be based on it. From this reasoning a current can be defined as the rate of change of charge. The greater the rate, the greater also is the current. The unit of charge is the coulomb. The charge on an electron is 1.602×10^{-19} coulombs. In terms of charge, an ampere of current is defined as one coulomb per second. Thus if 0.5C passes a given point in one second we have a current of 0.5A. Likewise if 2C passes a given point in one second we have a current of 2A. Knowing that a current is simply a directed movement of electrons we can now express an ampere in terms of the number of electrons passing a given point. Since the charge on an electron is 1.602×10^{-19} C and an ampere by definition is one coulomb per second, the number of electrons are found by taking the reciprocal of the charge on an electron. On calculation the answer is 6.25×10^{18}

electrons per second. This is indeed a very large number of electrons.

Voltage

This is a term we meet very often both in everyday life and in the sphere of electronics. Whenever we walk into an electrical store and ask for an electrical appliance, we soon discover that the operating voltage of the appliance needs specifying. When we construct an electronic circuit and wish to verify certain aspects of its performance we might have to measure voltage at selected points in the circuit. What exactly do we understand by the term voltage. Earlier on we saw that a directed movement of electrons within a material, required an electromotive force (e.m.f.). Electrons moving in a preferred direction cannot do so without encountering opposition. Atoms and other electrons are certainly going to be in the way, and collisions will occur. This opposition is in fact the resistance of the material. If the cross-sectional area in which the electrons have to move is small, we will naturally expect more collisions to occur. In other words the resistance for a small cross-sectional area will be greater than that for a large cross-sectional area. An e.m.f. acts to overcome the effect of resistance and make way for an electric current.

An e.m.f. is always associated with a potential difference (p.d.). The greater the e.m.f. the greater also is the potential difference between the terminals across which the e.m.f. is present. Potential difference is thus a measure of e.m.f. The term voltage is often used in place of potential difference and is intended to carry the same meaning. The unit of potential difference is the volt (V). The volt is defined as that voltage which appears across a resistance of one ohm when the current through it is one ampere.

Power

Like voltage the term power is also familiar to us. The power rating is an important quantity stated in most specifications relating to electrical equipment. Perhaps we can recollect several occasions when we have walked into an electrical store and the salesman has asked what wattage we are after. Now, the watt is the unit of power. It is referred to the more fundamental unit of energy, the joule. One watt is defined as one joule per second. The appearance of the 'per second' immediately indicates to us that it has something to do with rate. Well, it certainly has since power which is expressed in watts is defined as the rate of utilization of energy. The physical significance of the term power can be appreciated as follows. When current is forced to flow in a material having electrical resistance, the temperature of the material will rise. This occurs due to electrical energy being converted to heat energy within the material. The basic mechanism which brings about this energy conversion is seen to be the collisions between electrons, and between electrons and atoms in the material. If the material had no resistance whatsoever then no energy will be utilized and hence there is no requirement for the use of the term power. At this stage we could see intuitively that power in some way has to be related to voltage and current, since current through a sample of material, and the voltage across the points at which the current enters and leaves the material, are governed by the resistance. This is in fact the case; when we multiply voltage by current for a given situation we obtain power. We can state this in equation form as

$$P = VI \text{ watts} \quad \ldots \quad 4.1$$

A law of pivotal importance in electrical science and electronics due to Georg Simon Ohm, and known internationally as Ohm's Law can be invoked to obtain alternative expressions for power. Ohm's Law in essence serves to enlighten us on the relationship that exists between the current flowing in a resistance and the voltage across it. It states that the current is directly proportional to the voltage. If an increase or decrease in the value of a given variable quantity causes another to change in a similar manner we say that the variables are directly proportional to one another. A proportional relationship between two variables can be made an equality by introducing what is known as a constant. The constant in Ohm's Law is the resistance R. Ohm's Law can be very elegantly presented in mathematical form as

$$I = \frac{V}{R} \quad \text{.......... } 4.2$$

This equation may be rearranged as occasion demands. Useful alternative forms are

$$V = IR \quad \text{.......... } 4.3$$

and $$R = \frac{V}{I} \quad \text{.......... } 4.4$$

Alternative expressions for power using Ohm's Law are

$$P = \frac{V^2}{R} \quad \text{.......... } 4.5$$

and $$P = I^2 R \quad \text{.......... } 4.6$$

These expressions are obtained by substituting for I and V respectively in equation 4.1. In all the expressions above power is in watts, current is in amperes, voltage is in volts and resistance is in ohms.

The power level at which electronic circuits operate vary according to application. The output stage of an hi-fi amplifier will be capable of handling tens of watts, whereas its front end stages will typically operate at milliwatt (mW) levels. On the other hand CMOS circuits very much in vogue for digital applications operate at microwatt (μW) power levels.

In using formulae the general approach is to substitute values of all quantities in its basic form. That is if the quantity in question is current, the value should be expressed in amperes. For example a current of 0.1 mA is equivalent to 0.1×10^{-3} A. The same argument applies to resistance. A resistance of 1 KΩ should be expressed as $1 \times 10^3\,\Omega$. However considerable labour may be avoided by substituting values as they appear in practical situations, if we make an effort to remember the unit applicable to the final result for any given situation. This

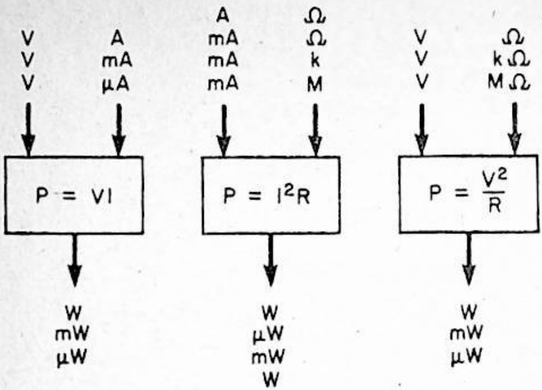


Fig. 4.3

information is presented pictorially in Figure 4.3, and applies to the equations for power. Each equation is viewed as a functional block which yields an output result when given input quantities are made available to it. In Figure 4.3b for instance, when the voltage and current are in volts and milliamperes, the power comes out in milliwatts.

Calculations involving power can be easily performed using a simple 4-function electronic pocket calculator, based on algebraic logic, and having a memory. In using a calculator the time taken for any given calculation can be reduced by resorting wherever possible to a chain calculation. Some expressions may be suited for such a calculation in its popular form while others may have to be rearranged. A chain calculation keeps the time taken to a minimum by keeping to a minimum the number of entries. Let us take as an example equation 4.5. This is a simple expression which requires no rearranging, and the sequence is that of a chain calculation whichever way the calculation is undertaken. Later we shall meet expressions of the type which require rearrangement.

Impedance

Earlier on we saw that the effect of resistance was to impede the flow of current. Inductance and capacitance can also behave in a likewise manner under certain conditions. The general term for this opposition is reactance, X. The opposition due to inductance is called inductive reactance, X_L, and that due to capacitance is called capacitive reactance, X_C. Like resistance reactance is also measured in ohms. Impedance is the term employed to describe the joint opposition to the flow of current caused by the presence of resistance and reactance.

The use of the term impedance is meaningful when the current and voltage waveforms in a circuit are sinusoidal. This is so since reactance is a function of frequency, and we may speak of frequency only when a quantity is varying with time in a cyclic manner. Inductive reactance is directly proportional to the value of inductance and frequency. That is it increases with inductance and frequency. On the other hand capacitive reactance is inversely proportional to the value of capacitance and frequency. That is, it decreases with capacitance and frequency. Thus the reactance of an inductor will be infinite at infinite frequency, while that of a capacitor will be zero at infinite frequency. The converse applies for zero frequency.

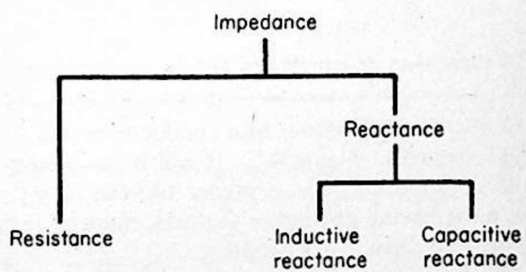


Fig. 4.4 Breakdown of impedance into its constituents.

The breakdown of impedance into resistance and reactance is illustrated in Figure 4.4, and serves to pictorially summarise the material covered above.

Admittance

The term admittance is intended to convey a meaning which is exactly opposite to that conveyed by impedance. It is a term employed to describe the ease with which current can flow. Like with impedance, admittance too can be separated into two parts, namely conductance and susceptance. By definition conductance is the reciprocal of resistance. The reciprocal of a number is the result one arrives at when unity is divided by that number. Thus the reciprocal of say, 5 is 0.2. By definition susceptance is the reciprocal of reactance. Admittance, conductance, and susceptance are measured in siemens, S.

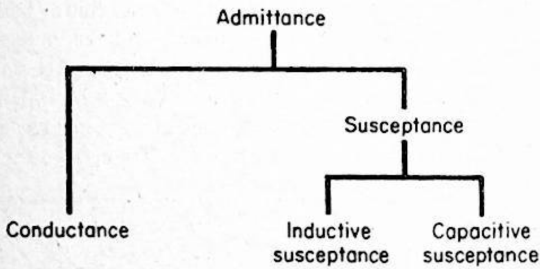


Fig. 4.5 Breakdown of admittance into its constituents.

The breakdown of admittance into conductance and susceptance is illustrated in Figure 4.5. It will be seen from the figure that like reactance, susceptance too can be separated into two parts having distinctive features, namely, inductive susceptance and capacitive susceptance.

Illustrative Examples

Example 4.1

A repetitive waveform has a period of 1 millisecond. What is its frequency.

The frequency is given by

$$f = \frac{1}{T}$$

Therefore

$$f = \frac{1}{1 \times 10^{-3}} \text{ Hz}$$

$$= 1000 \text{ Hz}$$

Example 4.2

A signal waveform has a frequency of 10 kHz. Find its period, and express the answer in microseconds.

The period is given by

$$T = \frac{1}{f}$$

Therefore

$$T = \frac{1}{10 \times 10^{3}} \text{ sec}$$

$$= 10^{-4} \text{ sec.}$$

To convert to microseconds we multiply the answer by 10^6
Thus

$$T = 10^{-4} \times 10^{6}$$

$$= 10^{2}\ \mu S$$

Example 4.3
The emitter resistor in a transistor amplifying stage is 2.7 k. The voltage across it is 2 V. Find the current flowing in the resistor.

This is an example where we must apply Ohm's Law. From equation 4.2

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

If the voltage is expressed in volts and the resistance is expressed in kilohms, the current is obtained in milliamperes. Substituting the given values

$$I = \frac{2}{2.7} \text{ mA}$$

$$= 0.74 \text{ mA}$$

Example 4.4
The voltage across a resistor and the current through it are measured and found to be 5.6 V and 1 mA respectively. Find the value of resistance.

Once again we use Ohm's Law. From equation 4.3b

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

If we substitute for voltage in volts and current in milliamps the resistance is obtained in kilohms. Thus

$$R = \frac{5.6}{1} \text{ k}\Omega$$

$$= 5.6 \text{ k}\Omega$$

Example 4.5
Find the power dissipated in the resistor of example 4.4. From equation 4.1 the power is given by

$$P = VI \text{ watts}$$

If we substitute for voltage in volts and the current in milliamps, the power comes out in milliwatts. Therefore

$$P = 5.6 \times 1 \text{ mW}$$

$$= 5.6 \text{ mW}$$

Example 4.6
The current flowing through a 2.7 k resistor is found to be 2.5 mA. Find the power lost in it. From equation 4.6 the power is given by

$$P = I^2 R \text{ watts}$$

If we substitute for current in milliamps and resistance in kilohms the power is obtained in milliwatts. Therefore

$$P = (2.5)^2 \times 2.7 \text{ mW}$$

$$= 16.875 \text{ mW.}$$

Chapter 5

LINEAR CIRCUIT COMPONENTS

Components used in an electronic circuit fall into two groups, namely linear and non-linear. A component is said to be linear if the voltage–current relationship is linear. An increase in one variable parameter should force the other to change in a proportionate manner. When the two variables are plotted on a linear graph paper, the result should be a straight line with a given slope.

Linear circuit components are sub-divided into passive and active components. A component is said to be passive if it does not deliver signal power to the surroundings. By surroundings we mean the circuit in which the said component resides. In terms of this definition, the resistor, capacitor, inductor, and transformer fall into this category. A component is said to be active if it delivers signal power to the surroundings. In terms of this definition, the thermionic valve, transistor, and the tunnel diode fall into this category. Both passive and active components find wide application in electronics and hence a good understanding of their behaviour from a qualitative viewpoint is of considerable importance to the hobbyist.

Passive Circuit Components

The Resistor

A resistor is a component that exhibits the property of resistance. It is used in circuits to perform a variety of functions, the most familiar being to limit a current, and to attenuate a voltage to a required value. A resistor absorbs electrical energy. Since energy cannot be created nor destroyed this energy appears as heat energy and causes the temperature of the resistor to rise. Physically, this energy conversion occurs due to electrons colliding with other electrons and atoms inside the resistor material. An increase in the energy conversion rate above a given limit will destroy the resistor since in such a situation the heat will not be transferred to the surrounding air, fast enough. The power rating of a resistor is a safeguard against this eventuality, and should never be exceeded under any circumstances.

Series Connection of Resistors

When two or more resistors are connected in series the total resistance is increased. The diagram in Figure 5.1 shows a

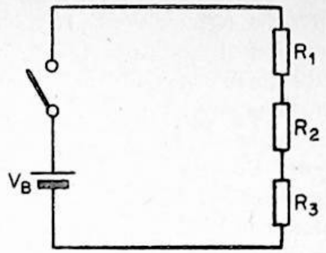


Fig. 5.1 Series connection of resistors.

circuit which consists of a battery, three resistors, and a switch connected in a series fashion. On closing the switch, electrons travel from the negative terminal of the battery to its positive terminal along the path provided by the series connection of the three resistors. This being the only path available, the same current has to flow through each resistor. Due to this reason the total opposition encountered by the current is represented by the algebraic sum of the resistances in the circuit. Therefore the total resistance, R_T, is given by

$$R_T = R_1 + R_2 + R_3 \quad \ldots\ldots\ldots\ldots \quad 5.1$$

This reasoning is valid for any number of resistors connected in series. If the resistances are equal in value, then

$$R_T = nR \quad \ldots\ldots\ldots\ldots \quad 5.2$$

where n is the total number of resistors in the series circuit, and R is the value of each resistance.

Parallel Connection of Resistors

When two or more resistors are connected in parallel the total resistance is reduced. Its value is always less than the smallest resistance value present in the combination. The diagram in

Figure 5.2 shows a circuit which consists of a battery, a switch, and three resistors connected in parallel. On closing the switch electrons are able to travel from the negative terminal to the positive terminal by three paths, provided by the parallel connection of the resistors. This current would divide between the three paths and the exact proportion of the total current along any given path will be governed by the resistance, of that path. The higher the resistance the smaller the current. Due to this reason the total resistance of a parallel connection is the algebraic sum of the reciprocal of each resistance in the combination. Therefore the total resistance is given by

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \quad \ldots\ldots\ldots\ldots \quad 5.3$$

This reasoning is valid for any number of resistors connected in parallel. If the resistances are equal in value, then

$$\frac{1}{R_T} = \frac{n}{R} \quad \ldots\ldots\ldots\ldots \quad 5.4$$

where n is the total number of resistors in the parallel combination and R is the value of each resistance.

Rearranging equation 5.4

$$R_T = \frac{R}{n} \quad \ldots\ldots\ldots\ldots \quad 5.5$$

In practical work we often wish to know the total resistance when two unequal resistors are connected in parallel. An expression for this is easily obtained from equation 5.3. If R_1 and R_2 are the resistances in question, then

$$R_T = \frac{R_1 R_2}{R_1 + R_2} \quad \ldots\ldots\ldots\ldots \quad 5.6$$

Sometimes the desired value of total resistance and the value of one resistor is known and we wish to compute the other. Suitable expressions for this purpose are obtained by solving equation 5.6 for R_1 and R_2 in turn. This gives

$$R_1 = \frac{R_T R_1}{R_2 - R_T}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.7$$

$$R_2 = \frac{R_T R_1}{R_1 - R_T}$$

Equation 5.3 must be used whenever the number of unequal resistors exceeds two.

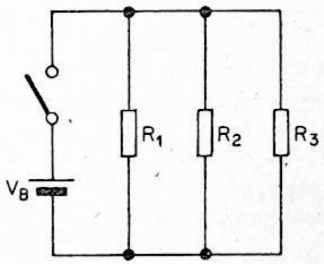


Fig. 5.2 Parallel connection of resistors.

Capacitor

A capacitor is a component which exhibits the property of capacitance. In its simplest form it consists of an insulating or dielectric material placed between two conducting plates. Unlike a resistor which absorbs electrical energy, a capacitor is able to store electrical energy. This energy is stored in the form of an electric field. An electric field may be imagined as a region of space where an unit electric charge experiences an electric force. Essentially the field is associated with the charge deposited on each plate of the capacitor. Capacitors are available in various types. The choice of a given type depends on the application. The principal types available and their recommended application areas are given in Figure 5.3 and should serve as a quick reference. When speaking of

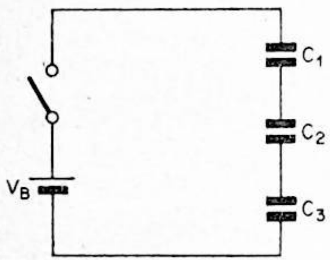


Fig. 5.3 Series connection of capacitors.

capacitors three important quantities appear very often. These are charge, (Q), capacitance, (C), and working voltage, (V_W). The charge, Q, is related to the capacitance, C, and the voltage, V, by the expression

$$Q = VC \text{ coulombs} \quad \ldots\ldots\ldots\ldots\ldots \quad 5.8$$

Some idea of the size of a capacitor may be had by considering the product of capacitance and working voltage. The larger this product is, the larger we may expect the capacitor to be. Since the construction technique, and the permittivity of the dielectric material influences the size as well, the above argument holds true for a given family, and is a point to watch.

Series Connection of Capacitors

A desired value of capacitance may be obtained by connecting capacitors in series. A series connection gives a capacitance value less than the lowest value of capacitor in the series circuit. The expressions for the series connection are similar in form to those for the parallel connection of resistors. This fact is useful in remembering the expressions. The similarity arises due to the manner in which the charge distributes itself when a series connection is made. We find that an equal charge is carried by each capacitor. When resistors are connected in parallel we have already seen that an equal voltage is present across each resistor. The reasons for the similarity become clear when we consider these observations in conjunction with the formulae for the charge in a capacitor, and the voltage across a resistor.

To keep the treatment consistent with that for resistors, we shall consider the state of affairs when three capacitors are connected in series via a switch to a battery, as shown in Figure 5.3. When the switch is closed an equal charge is acquired by each capacitor. Since the battery voltage equals the algebraic sum of the voltages across each capacitor, and these voltages are obtained when we divide the charges by the respective capacitances, we arrive at the expression given below for the total capacitance:

$$\frac{1}{C_T} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.9$$

When the number of capacitors is limited to two, the total capacitance is given by:

$$C_T = \frac{C_1 C_2}{C_1 + C_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.10$$

Stated in words the total capacitance when two capacitors are connected in series is given by the product of the individual capacitances divided by the sum of the individual capacitances. When the number of capacitors connected in series is n and have equal capacitance, then the total capacitance is given by:

$$C_T = \frac{C}{n} \qquad \text{5.11}$$

Parallel Connection of Capacitors

A desired value of capacitance may also be obtained by connecting capacitors in parallel. A parallel connection gives a total capacitance which is the algebraic sum of the individual capacitances. The expression for the parallel connection are similar in form to those obtained for the series connection of resistors. This fact is useful in remembering the expression. This similarity arises due to the voltages being equal in the case of the capacitor, and the current being equal in the case of the resistor. These observations taken in conjunction with the formulae for the charge in a capacitor and the current through a resistor accounts for the similarity.

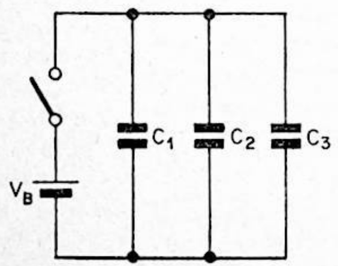


Fig. 5.4 Parallel connection of capacitors.

Let us now consider the circuit shown in Figure 5.4. On closing the switch, charge is transferred to each capacitor in proportion to its capacitance. Therefore in the parallel case the charges acquired by each capacitor are unequal. However since the voltage across each capacitor is equal and the total charge is the algebraic sum of the individual charges, we obtain the following expression for the total capacitance:

$$C_T = C_1 + C_2 + C_3 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 5.12$$

When n capacitors of equal value are connected in parallel, the total capacitance is given by

$$C_T = nC \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.13$$

Inductor

An inductor is a component which exhibits the property of inductance. In its simplest form it consists of several turns of wire. This structure is popularly referred to as a coil. The value of inductance depends on the number of turns and on the type of material placed along the axis of the coil. A magnetic property of the material, identified by the term permeability determines the influence it has on the inductance. The higher the permeability the higher the inductance for a given number of turns. By virtue of its inductance an inductor can store electrical energy. This energy is stored in the magnetic field created by the flow of current. Like resistors and capacitors, inductors too are widely used in diverse applications. Their use is almost a necessity in selective circuits operating at radio frequencies.

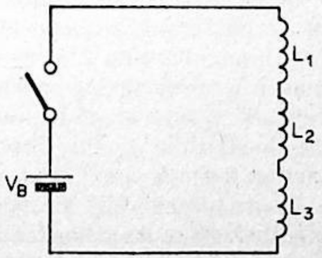


Fig. 5.5 Series connection of inductors.

Series Connection of Inductors

The total inductance resulting from a series connection depends on the manner in which the connection is made. If the connection is such that there is no magnetic coupling between them, then an expression for the total inductance is easily obtained. The diagram in Figure 5.5 shows the series connection of three inductors to a battery via a switch. It is assumed that there is no magnetic coupling between them. No sooner the switch is closed the current starts to rise from zero and continues to do so with a constant slope governed by the total inductance. Since no resistance is present, the current will go on increasing and reach infinity at infinite time. Now the current through an inductor cannot change without giving rise to what is known as a back e.m.f. The magnitude of this back e.m.f. depends on the rate of change of current and inductance. Since the same current flows through both inductors and the total backe.m.f. is the sum of the individual back e.m.f.s, the total inductance is given by the algebraic sum of the individual inductances. Therefore

$$L_T = L_1 + L_2 + L_3 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 5.14$$

The above expression does not hold if there is mutual coupling between the inductors. The effect of such coupling may be understood by considering just two inductors. The magnetic field produced by the current in one inductor influences the field associated with the other, and vice versa. This modifies the expression for the total inductance by adding or subtracting from it a term having the value 2M. M is the mutual inductance, and quantifies the coupling present between the two inductors. Now, the inductors can be connected in series in two ways, depending on the winding disposition. If the connection is such that the back e.m.f. is greater than that due to the two inductors in a magnetically isolated situation, they are said to be connected in a series aiding fashion, and the sign assigned to the mutual inductance term is positive. If the back e.m.f. is less, then the sign is negative, and the inductors are said to be connected in a series opposing fashion.

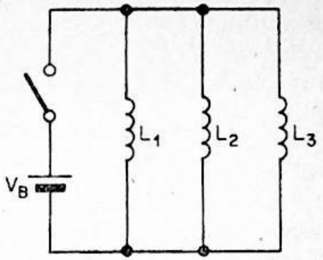


Fig. 5.6 Parallel connection of inductors.

Parallel Connection of Inductors

The diagram in Figure 5.6 shows the parallel connection of three inductors to a battery via a switch. If there is no magnetic coupling between them the total inductance is given by

$$\frac{1}{L_T} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.15$$

As to why the expression takes this form is easily appreciated by considering the behaviour from a back e.m.f. point of view. No sooner the switch is closed, a current is set up in the inductor, and starts to rise from zero at a fixed rate governed by the inductance and supply voltage. Since the inductances differ the rate of increase in each case would also differ. However the back e.m.f.s are the same for all the inductors and equal the supply voltage for the situation depicted in the figure. This reasoning together with the fact that the current in each inductor is inversely proportional to the inductance leads to the expression given. It is worthwhile noting that this expression is similar in form to that for the parallel connection of resistors.

As before we shall consider the case with mutual coupling by taking just two inductors. The magnetic field associated with one will influence the other, and vice versa, and this would modify our expression. If the mutual inductance is M, the total inductance is given by

$$L_T = \frac{L_1 L_2 - M^2}{L_1 + L_2 \pm 2M} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 5.16$$

The sign in the denominator will be positive if the inductors are connected in an opposing fashion. It will be negative if the inductors are connected in an aiding fashion. The connection is said to be aiding or opposing depending on whether the back e.m.f. is greater or less than the value associated with the magnetically isolated situation.

Transformer

When the magnetic field associated with one current carrying circuit undergoes a change, a current appears in any other circuit which comes within the confines of this field. The transformer is a passive circuit component which exploits this phenomenon. It basically consists of two windings, called the primary and secondary. When the current in the primary winding is forced to fluctuate by applying a time-varying voltage to it, the magnetic field due to this current also fluctuates. This causes a voltage to appear across the terminals of the secondary winding. The magnitude of this secondary voltage will depend on the turns-ratio and magnetic coupling between the two windings. If we terminate the secondary winding by a load circuit, a fluctuating secondary current will flow, giving rise to a fluctuating magnetic field. This field will interact with the original field causing some current readjustment in the primary winding. The concept of reflected impedance arises due to this effect.

An ideal transformer will transfer power from the primary side to the secondary side without a power loss. In practice the resistance of the copper wire employed for the winding and the iron losses in the core material will cause the secondary power to be less. The efficiency of the transformer expressed as a percentage indicates to us how well the power is transferred from the primary to the secondary. For an iron cored transformer, typical efficiencies are in the region of 70–90%.

The transformer is a very useful circuit component. By a suitable choice of turns-ratio it could be employed to step-up or step-down a voltage or current. A step-up in voltage is always accompanied by a step-down in current and vice versa. This has to be necessarily so since the transformer is a device which cannot give power gain. Another application of the transformer is in matching. The resistance looking into a given pair of terminals may be too high for a specific need. By using a transformer it is possible to transform this resistance to the required value. If a transformer is connected to a load resistance in the manner shown in Figure 5.7,

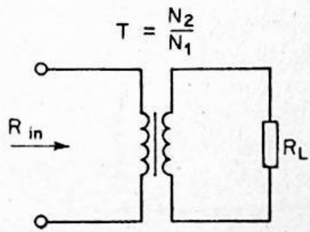


Fig. 5.7 Resistance matching using a transformer.

the ac resistance seen looking into the primary is given by

$$R_{in} = \frac{R_L}{T^2} \qquad \text{5.17}$$

In this equation T is the turns-ratio. T is defined as the ratio of the number of secondary turns (N_2) to the number of primary turns (N_1). Equation 5.17 reveals that R_{in} is greater than R_L if T is less than unity. This condition is satisfied when the number of secondary turns is less than the number of primary turns. It is smaller than R_L if T is greater than unity, and equal to R_L if T is equal to unity. These observations show that by a suitable choice of turns-ratio we are able to match any two resistances. It must be emphasized that the transformer is an ac device and can be used for resistance matching, only in ac situations.

Active Components

The principal components in this category are the thermionic valve, the transistor and the tunnel diode. We shall not undertake a discussion of the valve due to its diminished importance as an active device. The transistor will be subjected to a detailed treatment. We shall exclude treatment of the tunnel diode since it is of limited interest to the hobbyist.

The Transistor

The transistor is a versatile three-terminal semiconductor device which can be called upon to perform a variety of circuit functions as occasion demands. Transistors are broadly classified as being either bipolar or unipolar. A bipolar transistor is one whose operation depends on both types of charge carrier, namely the electron and the hole. The unipolar transistor relies only on one type.

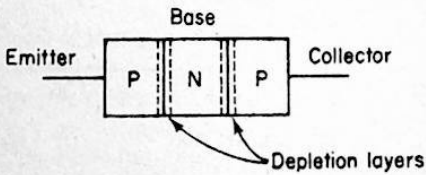


Fig. 5.8 Basic structure of pnp transistor.

Junction Transistor

The junction transistor is the most important device in the bipolar group. The basic structure of a p-n-p device is shown in Figure 5.8. It consists of a thin slice of n-type material sandwiched between two slices of p-type material. When n-type and p-type material are brought together what is known as a depletion layer is formed on either side of the junction. This is a region which is depleted of charge carriers. Since a transistor has two junctions we would expect to encounter two depletion layers, and this is in fact so. In putting the transistor to use, it is necessary to provide it with dc potentials. This is known as biasing. It is possible to either forward bias or reverse bias a pn junction. The effect in each case would be to modify the width of the depltion layer. Forward bias decreases it and reverse bias increases it. A forward-biased and a reverse-biased pn junction are shown in Figure 5.9. In the forward biased case the positive and negative terminals of

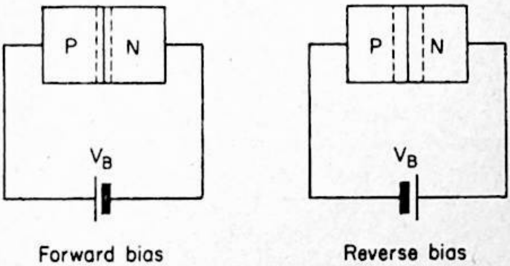


Fig. 5.9 Forward and reverse biased pn junctions.

the battery are connected respectively to the p-type and n-type material. In the reverse biased case the positive and negative terminals are connected respectively to the n-type and p-type material. Biasing determines the current flow across a junction. Forward bias gives rise to a current, while in the ideal case reverse bias does not. However due to thermal effects and material imperfections an extremely small current flows in a reverse biased junction. This current is called a leakage current.

For a transistor to function as linear device we must ensure that its junctions are suitably biased. The base-emitter junction must be forward biased, and the base-collector junction must be reverse biased. The question as to how current flow takes place across a reverse biased junction, is bound to arise in our minds. This is what transistor action is all about and we shall soon discover the answer when discussing the operating principle of the device.

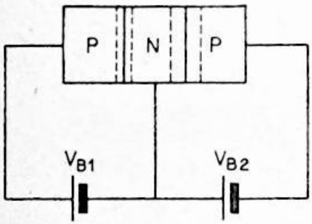


Fig. 5.10 pnp transistor biased for linear operation.

The diagram in Figure 5.10 shows a pnp transistor biased for operation as a linear device. The base-emitter junction is forward biased, and the base-collector junction is reverse biased. Its operation is as follows. Holes crossing over to the base region from the emitter come under the influence of an accelerating field created due to the reverse biasing of the base-collector junction. Therefore they are quickly swept across the base-collector junction and give rise to an electron current in the collector-emitter external circuit. All the holes that enter the base region do not find their way into the collector region. Some recombine with electrons which are the majority carriers in n-type material. For each hole that recombines, an electron enters the base region from the external base circuit. These events cause a small electron current to flow into the base terminal. It must be emphasized that the concept of a hole as a current carrier is meaningful only for events within a semiconductor material. The current in the external circuit is always associated only with electrons. The directions of circuit symbols current flow for a pnp and npn transistor are shown in Figure 5.11. The emitter current is simply the algebraic sum of the base and collector currents. For a npn transistor the relationships are similar, except that now the battery polarities and current directions are reversed.

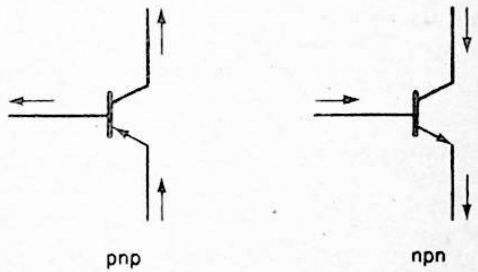


Fig. 5.11 Circuit symbols and direction of current flow for pnp and npn transistors.

There are three ways in which the transistor can be put into service as an active device. These are identified by the transistor terminal which we select to be common to the input and output circuits. Modes of operation made possible on this basis are the common-base, common-emitter, and common-collector. Sometimes the terms grounded-base, grounded-emitter, and grounded collector are used instead and conveys exactly the same meaning.

The transistor is often called upon to amplify signals. The current gain as stated by the manufacturer indicates to us, how well the device, can perform this function. Transistor specifications usually give the common emitter current gain, h_{fe}. This is defined as the ratio of a change in collector current to a change in base current, at a given value of collector current.

The common emitter current gain is a very useful quantity. A knowledge of it enables us to calculate the collector current, for a given base current, and the input resistance. If the collector current is known using h_{fe}, we can calculate the base current.

The collector current is given by

$$I_C = I_B h_{fe} + I_{co}(1 + h_{fe}) \quad \ldots\ldots\ldots\ldots\ldots \quad 5.18$$

where I_B is the base current and I_{co} is the leakage current.

For a silicon transistor I_{co} is very small and can often be neglected.

Therefore

$$I_C = I_B h_{fe} \quad \ldots\ldots\ldots\ldots\ldots \quad 5.19$$

Rearranging

$$I_B = \frac{I_C}{h_{fe}} \quad \ldots\ldots\ldots\ldots\ldots \quad 5.20$$

The input resistance for the common-emitter and common-collector modes is given by

$$R_{in} = h_{fe}(R + r_e) \quad \ldots\ldots\ldots\ldots\ldots \quad 5.21$$

where r_e is the resistance of the forward-biased emitter-base junction, and R is the total resistance of the external emitter circuit.

The resistance r_e depends upon the emitter current, and is given by

$$r_e = \frac{25}{I_E} \quad \ldots\ldots\ldots\ldots\ldots \quad 5.22$$

where I_E is the emitter curren in milliamperes.

Unipolar Transistor

We have seen that the conduction processes taking place in the junction transistor make use of both types of charge carrier. The unipolar transistor in contrast makes use of one type to carry most of the current. Unipolar transistors are better known as field-effect transistors. There are two important types. These are the junction field-effect transistor (JFET) and the metal oxide semiconductor field-effect transistor (MOSFET). In this chapter we shall only consider the JFET due to its importance in linear applications. The MOSFET which is of value as a switching device will be treated in the next chapter.

The JFET has three electrodes. These are known as the source, gate, and drain, and correspond to the emitter, base, and collector of a bipolar device. The conduction of current between the source and drain can be controlled by varying the voltage of the gate with respect to the source. The diagram in Figure 5.12 shows the basic structure and circuit

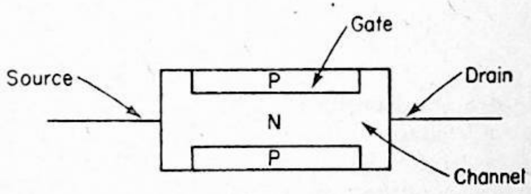


Fig. 5.12 Basic structure of JFET.

symbols of a JFET, and serves to explain its principle of operation. Semiconductor material of either n or p type is surrounded lengthwise by p or n type material respectively. The material surrounded in this manner is known as the channel. To each end of the channel electrodes are attached and these are known as the drain and source. A depletion layer is formed on either side of the metallurgical junction resulting from the outer material making contact with the channel. An electrode is attached to the outer layer and this electrode is known as the gate. In normal operation, the gate to source junction is reverse biased. As a result the depletion layer penetrates further into the channel and interferes with the flow of current in the channel. In this way the gate exercises control over the source to drain current. The input impedance at the gate electrode is extremely high, since it is reverse biased with respect to the source. The only current which flows in the gate circuit is the leakage current, and this is extremely small in silicon devices. The conduction between the source and drain is entirely due to majority carriers, in the channel. Thus for a n-channel JFET, conduction is due to electrons, and for a p-channel JFET, conduction is due to holes.

The relationship between the drain current, I_D and the gate to source voltage, V_{GS}, is determined by the process employed in the manufacture of the device. If the junction has been obtained by a diffusion process, the drain current is given by

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 \qquad \ldots\ldots\ldots\ldots\ldots \quad 5.23$$

where I_{DSS} is the drain current when V_{GS} is zero, and V_P is the pinch off voltage.

The significance of the term pinch-off voltage may be understood by considering the manner in which the drain current increases with drain-source voltage, while the gate-source voltage is held at zero volts. As V_{DS} is increased from zero volts, the drain current too commences to increase from zero. At some point of V_{DS}, I_D flattens off and does not respond to any further changes in V_{DS}. The value of V_{DS} at which this happens is referred to as the pinch-off voltage, V_P. By reverse biasing the gate-source junction by a voltage equal to the pinch-off voltage, we are able to reduce I_D to zero, and this argument forms the basis for an alternative definition of the pinch-off voltage. In the above explanation we held the gate-source voltage at zero. We could very well have held it at some other convenient voltage of either polarity. This would in effect have given us different values for V_P and I_D.

Linear Integrated Circuits

The integrated circuit is now well established as a circuit element. Circuit functions which were traditionally performed by several passive and active discrete elements are now taken over by integrated devices. We have already seen the basic structure of bipolar and unipolar transistors. We know that these consist of n-type and p-type semiconductor materials. We also know that these materials when in contact give rise to a junction which has a profound influence on device operation. Now an integrated circuit is one which has several such junctions in one piece of semiconductor material. Some behave as transistors and diodes, while others function to isolate sections of the circuit as is necessary, by exploiting the properties of a reverse biased junction. Resistors are incorporated by controlling the concentration of charge carriers in the material. This is a very simplified view of the integrated circuit. The design and manufacture of integrated circuits is a highly specialized field, and very elegant technology and concepts are associated with the processes involved.

A linear integrated circuit is one which can be used in a linear mode of operation. The operational amplifier and the voltage regulator integrated circuits are examples of linear devices.

Illustrative Examples

Example 5.1

Three resistors whose values are 120, 180, and 220, are connected in series. Find the total resistance.

Solution

The total resistance is given by equation 5.1. Therefore

$$R_T = R_1 + R_2 + R_3$$

Substituting the given values

$$R_T = 120 + 180 + 220$$

$$= 520 \text{ ohms.}$$

Example 5.2

Three resistors having values 120, 180 and 220 are connected in parallel. Find the total resistance.

Using equation 5.3, the total resistance is given by

$$\frac{1}{R_T} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

Substituting the values given

$$\frac{1}{R_T} = \frac{1}{120} + \frac{1}{180} + \frac{1}{220}$$

$$= \frac{1}{1.2 \times 10^2} + \frac{1}{1.8 \times 10^2} + \frac{1}{2.2 \times 10^2} \quad \text{. . . Step 1}$$

$$= \frac{1}{10^2} \quad \frac{1}{1.2} + \frac{1}{1.8} + \frac{1}{2.2}$$

$$= \frac{1}{10^2} (0.83 + 0.55 + 0.44) = \frac{1.83}{10^2}$$

$$\therefore R_T = \frac{10^2}{1.83} = 54.5\Omega$$

Comments: In handling calculations involving reciprocal terms it is recommended that the denominator of each term is expressed as indicated in step 1. This technique will make an error less likely.

Example 5.3

Two resistors having values 39Ω, and 47Ω, are connected in parallel. Find the total resistance.

Using equation 5.6

$$R_T = \frac{R_1 R_2}{R_1 + R_2}$$

Substituting values

$$R_T = \frac{39 \times 47}{39 + 47}$$

$$= 21.3\Omega$$

Chapter 6

NON-LINEAR CIRCUIT COMPONENTS

A non-linear circuit component is one which exhibits a voltage-current relationship, which is either not linear or has a sharp discontinuity. In other words, equal changes in one variable, say the voltage, will not produce equal changes in the other. Graphical representation of such behaviour will produce either a curve or a sharp discontinuity, instead of a straight line with a given finite slope. Since the most common application of a non-linear device is in a switching situation, these devices are often referred to as switching elements. The principal components of interest to the hobbyist in this category are the general-purpose diode, zener diode, thyristor, triac and the digital integrated circuit. Both bipolar transistors, and unipolar transistors of the junction-gate variety, can be made to behave as switching elements.

The Junction Diode

The basic structure of the junction diode together with its circuit symbol are shown in Figure 6.1. It consists of n-type and p-type semiconductor material. The two types of material are in close contact and forms a pn junction. A depletion layer is formed on either side of the junction. In biasing the diode we modify the width of this layer.

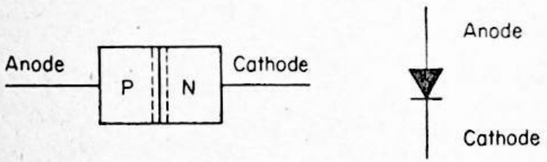


Fig. 6.1 The basic structure and symbol of junction diode.

A junction diode can be either forward-basied or reverse-biased. In the forward-biased condition its resistance is low and a forward current, limited by the total series resistance flows in it. In the reverse-biased condition, its resistance is high and a very small reverse-current flows. The diagrams in Figure 6.2 show the two ways in which the junction diode can be biased. Every diode has a current rating which must not be exceeded. In Figure 6.2a a series resistor is included to limit the current to a safe value. If this resistance is high in comparison to the forward resistance of the diode, the battery voltage will be dropped almost entirely across it. In such circumstances the diode current is simply given by

$$I = \frac{V}{R} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 6.1$$

The forward voltage drop across a germanium diode is typically 0.3 V, and that of a silicon diode is 0.6 V. This is an useful fact to keep in mind whenever confronted with forward biased pn junctions.

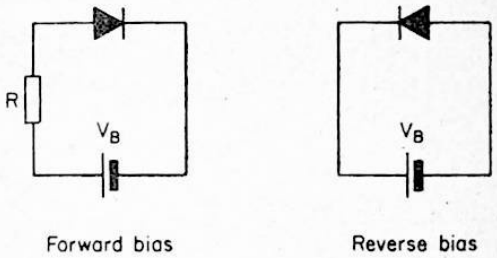


Fig. 6.2 Forward and reverse biasing a junction diode.

The maximum reverse voltage that we may apply across a junction-diode is limited by its design. The manufacturer gives a rating which should be watched in any diode application. Exceeding the value will cause a sudden increase in the small reverse current, and unless this is limited by an external resistance, the diode will suffer damage.

The voltage-current relationship for a junction diode is shown in Figure 6.3. This diagram shows graphically all the features we have discussed so far.

Up to now we have avoided any reference to the mechanism of operation. This is easily understood by making use of the knowledge we have already acquired concerning atomic structure and electrical conduction in solids. In the forward-biased condition, holes which are the majority carriers in the p-type material cross over to the n-type material where they assume the status of excess minority carriers. Likewise electrons

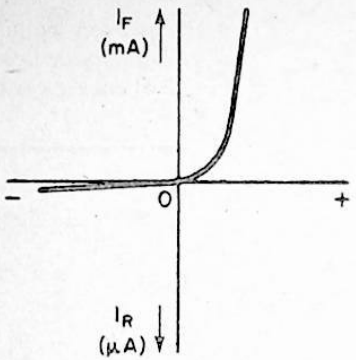


Fig. 6.3 Voltage vs Current characteristic of typical junction diode.

which are majority carriers in the n-type material cross over to the p-type material, and assume a similar status. The unbalance in the minority carrier equilibrium brought about by this migration causes each type of carrier to diffuse away from the edge of the depletion layer and travel towards the terminal contacts. Being immersed in a sea of majority carriers, these minority carriers recombine with them. As a result of recombination, for each hole that enters the n-type material, and electron is drawn into the material from the external circuit, and for each electron that enters the p-type material an electron is expelled to the external circuit. These events give rise to a forward diode current. By the use of the term recombination we signify an event where an electron fills a vacancy in a donor atom. By definition a vacancy for an electron is a hole.

An alternative argument may be pursued as follows. At any p-n junction, what is known as a barrier potential exists. This potential acts to oppose current flow across the junction. When we forward bias the junction, we are in fact reducing the barrier potential, and permitting a current to flow. A decrease in the barrier potential is accompanied by a decrease in the width of the depletion layer.

In the reverse-biased condition, the migration of majority carriers is prohibited. An extremely small leakage current is observed in this condition, and this is due to the thermal generation of hole-electron pairs and surface effects. In terms of our alternative argument, reverse biasing, in effect, increases the barrier potential. An increase in the barrier potential is accompanied by an increase in the width of the depletion layer.

The Zener Diode

This is a device whose operation depends on zener breakdown. When heavily doped semiconductor materials are used for a pn junction, the resulting depletion layer is very thin in width. Consequently the electric field across it is very large, and the junction breaks down on applying a very small reverse voltage. The exact value of breakdown voltage, V_Z, is controlled, by controlling the doping levels during manufacture. Breakdown of a pn junction in this manner, is known as Zener breakdown.

A junction may also breakdown due to ionization caused by fast moving high energy electrons. This type of breakdown is known as avalanche breakdown. Junctions which breakdown at reverse voltages less than approximately 5 V, do so due to the Zener effect. The avalanche effect predominates at higher voltages.

The voltage-current characteristic and the circuit symbol for a Zener diode are shown in Figure 6.4. A principal application of the Zener diode, is as a voltage reference element in power supply units. In this application as well as in many others the device is operated in the breakdown region of the characteristic. The salient feature about this region is that the current through the device can change over wide limits with hardly any change in the breakdown voltage, V_B. It is always good practice to establish the operating point well away from the knee of the characteristic to take full advantage of this

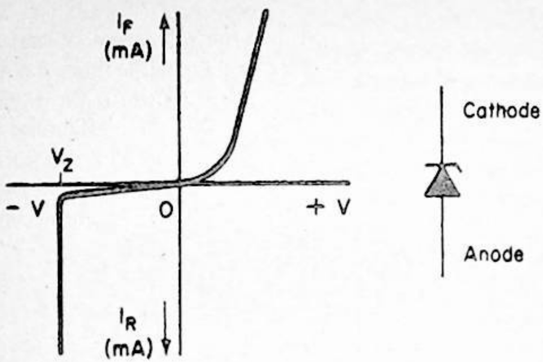


Fig. 6.4 Voltage vs Current characteristic of Zener diode and its circuit symbol.

feature. Zener diodes are classified according to power rating, and are available in a wide range of voltages. All voltage reference diodes are popularly called Zener diodes, even though the avalanche effect is responsible for the breakdown in devices having a reference voltage greater than about 8 volts.

The Light Emitting Diode

The light emitting diode or LED as it is popularly called is a semiconductor device which is capable of producing a light when a current is passed through it. The principal materials used in its manufacture are Gallium, Arsenic and Phosphorous. Gallium is an element which belongs to Group 3 of the Periodic Table of elements. Arsenic and Phosphorous are elements which belong to Group 5 of the Periodic Table.

The LED operates in the following way. Being a junction device it can be either forward or reverse biased. When forward biased electrons and holes recombine and in doing so produce photons. A photon may be regarded as a quantum or portion of light energy. It must be emphasized that this action of producing photons applies to pn junctions made from materials referred to above and does not concern junctions made from silicon or germanium. The colour of the emitted light depends on the composition of the constituent materials, and the luminous intensity depends on the forward current flowing across the junction. Luminous Intensity is expressed in millicandelas (mC). A typical red LED will produce a luminous intensity of 0.2 mC to 3 mC over a current range of about 10 mA to a 100 mA.

There are several points we must keep in mind when using a LED. First and foremost we must note that the forward voltage drop of a LED differs significantly from a normal junction diode. This voltage is a function of the forward current for any given type of LED, and can have a value between 1 and 2 V. In calculations a good rule of thumb would be to assume the forward voltage drop to be 1.8 V. The maximum ratings which apply to the forward current and peak inverse voltage should always be verified from the manufacturer's data. For most applications it will be found that a forward current of 10 mA gives an adequate luminous intensity. Since the forward voltage drop is much lower than the rail voltages normally encountered in equipment, it is necessary to use a resistor in series with the LED to drop the extra voltage. The value of this resistor is easily determined by an application of Ohm's Law. If the rail voltage is V_s and the LED forward voltage is V_f, the value of resistance required to operate the device at a current of I mA is given by dividing the difference of the two voltages, $V_s - V_f$, by the current I.

The Thyristor

The thyristor or silicon-controlled rectifier (SCR) as it is sometimes called, is a 4-layer, 3-terminal semiconductor device. It is the semiconductor counterpart of the thyratron. The thyratron is a thermionic gas-filled valve. It has been replaced by the thyristor in almost every application and is of interest only from a historical viewpoint. The principal use for the thyristor is in power control. The several salient features it possesses has opened up exciting new possibilities in the area of power control.

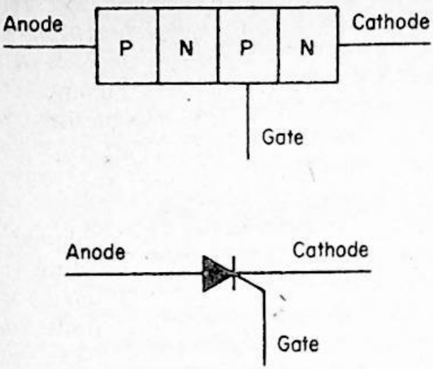


Fig. 6.5 Basic structure of thyristor and its circuit symbol.

The basic structure of a thyristor and the circuit symbol are shown in Figure 6.5. The electrodes are as identified in the figure. The thyristor like the diode behaves as a switch. A facility for switching on the device by applying a gate voltage or a specified anode voltage constitutes the basic difference between them. Once the device has been switched on, the current has to be reduced to below the so-called holding current, or the anode voltage removed, in order to switch it off. The gate loses control no sooner the device is switched into the on condition.

The thyristor is very useful in ac power control. By defining the conduction time over each cycle of the mains voltage, the power delivered to the load, can be controlled. This method of power control avoids the power loss associated with other methods such as employing series resistance.

The Triac

We have already seen that a thyristor is a device that can conduct in one direction. In power control applications a device which can conduct in either direction is of great value. The triac meets this requirement. By application of suitable gate voltages, the device may be made to conduct during the positive or negative half cycle of the input alternating voltage. The power transferred to the load is controlled by defining the fraction of time over which the device is allowed to conduct.

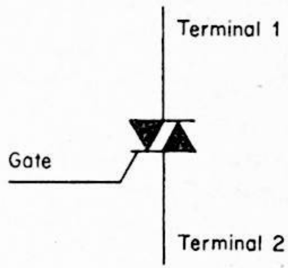


Fig. 6.6 The circuit symbol for a triac.

The basic structure of the triac may be imagined as consisting of two thyristors, one inverted with respect to the other, and placed alongside each other. The circuit symbol is shown in Figure 6.6.

Digital Integrated Circuits

A digital integrated circuit is a circuit element which operates in a non-linear mode. The inputs to a digital chip respond to logical signals. A logic signal is one to which a truth value is attached. A statement can be either true or false. A true statement can be represented by a high voltage and a false statement can be represented by a low voltage. These are identified by the logical symbols 1 and 0. Digital circuits employ two types of logic. These are known as combinational logic and sequential logic. Combinational logic may be performed using gates. The four main types of gates are the OR, NOR, AND, and NAND. These gates are available in integrated circuit form, and can be used to design complete systems. Sequential circuits are those which have a memory capability. Sequential logic is to do with such circuits and systems. Again these functions are available in integrated circuit form.

Digital integrated circuits are classified according to the technology employed, and are identified as a family. The most popular families in use today are TTL and CMOS. TTL stands for Transistor-Transistor-Logic, and CMOS stands for Complementary metal-oxide semiconductor. TTL is based on bipolar technology, where CMOS employs unipolar technology. A large range of logic functions are available from both families, and the choice depends on the application.

The Mosfet

It was mentioned earlier that unipolar transistors are available in two types, namely the JFET and the MOSFET. While both these devices rely on one type of charge carrier for conduction, there are fundamental differences in their structure and mechanism of operation. The JFET as its name implies depends very much on the normal junction behaviour. The channel conduction is controlled by controlling the depth to which the depletion layer penetrates into the channel. Conduction in a MOSFET on the other hand depends on the

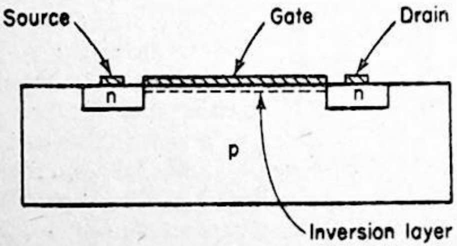


Fig. 6.7 Cross-sectional view of n-channel mosfet.

presence of a so-called inversion layer of charge carriers. A cross-sectional view of the basic structure is shown in Figure 6.7. On to one surface of a p-type substrate n-type regions are diffused. These n-regions behave as the source and drain. Conduction between the source and drain is made possible by an inversion layer which is produced only on application of a suitable potential to the gate electrode. The gate consists of a metal electrode placed on the top surface and isolated from it by an insulating oxide layer. The metal electrode, the insulating layer and the substrate behave as a parallel-plate capacitor. If a voltage positive with respect to the source is applied to the gate, a negative layer will appear on the substrate material in contact with the insulating layer on the opposite side. This layer is called an inversion layer. It functions to provide a conduction path between the source and drain. With no voltage applied to the gate electrode, the resistance between the source and drain is very high since the associated pn junctions are in effect reverse biased.

The input resistance of a MOSFET is very high. This is largely due to the presence of the insulating layer between the gate and source. The MOSFET is useful in switching applications.

Illustrative Examples

Example 6.1
A silicon junction diode is connected across a 15 V supply in a forward biased fashion. A 15 k resistor is used in series with the diode to limit the forward current. Estimate the current in the circuit.

Solution
Since the supply voltage is very much larger than the typical voltage drop of 0.6 V for a silicon diode, the current will be largely defined by the series resistance. Using equation 6.1

$$I = \frac{V}{R} = \frac{15}{15\,k} = 1\,mA$$

Example 6.2
It is required to operate a 3.3 V Zener diode from a 15 V supply at a current of 2 mA. Calculate the value of resistance that should be connected in series with the diode.

Solution
The value of resistance used should be able to absorb the voltage difference between the Zener voltage and the supply voltage. Using Ohm's Law

$$R = \frac{V_B - V_Z}{I}$$

$$= \frac{15 - 3.3}{2\,\text{mA}} = \frac{11.7}{2}$$

$$= 5.85\ \text{k}$$

In a practical situation the preferred value 5.6 k will be used.

Chapter 7

DC AND AC CIRCUITS

A dc circuit is one which is energised or excited by a dc voltage. There are several circuits which we must consider under this category. These are shown in Figure 7.1. However, before undertaking a discussion of these circuits it is instructive to take a quick look at two very important circuit laws. These laws are known as Kirchoff's Laws, and describe the basis on which the voltage is distributed in a series circuit, and the current is distributed at a junction.

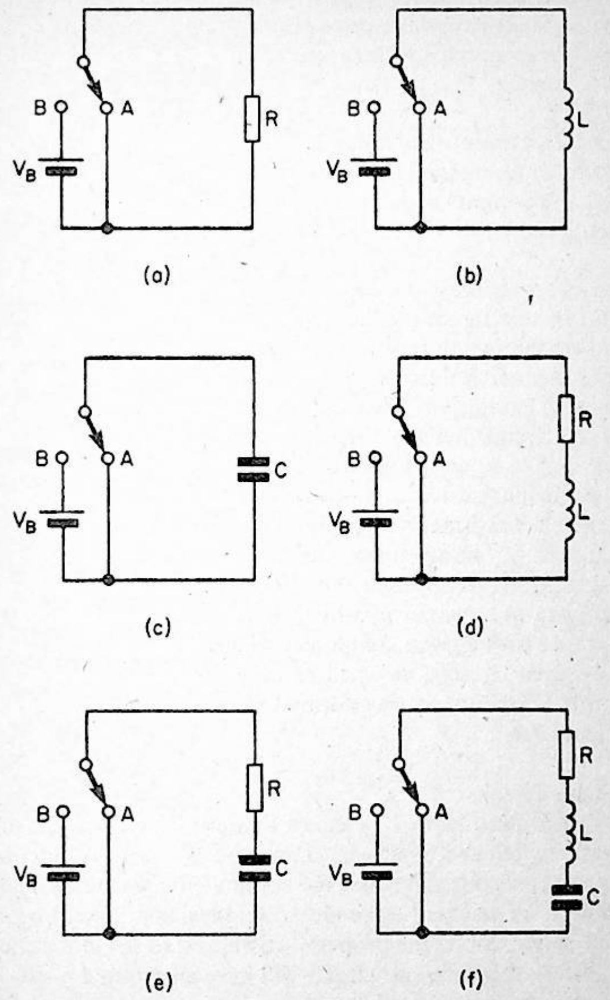


Fig. 7.1 DC circuits.

Kirchoff's voltage law states that the source voltage in any series circuit is equal to the algebraic sum of the voltage drops associated with each resistance. The application of this law is quite straightforward provided we remember that the voltage drop across a resistor is given by the product of current and resistance.

Kirchoff's current law states that the algebraic sum of the currents entering and leaving a junction is zero. In other words the currents which enter a junction equal the currents which leave.

Let us now return to the circuits shown in Figure 7.1. Each circuit in this figure can be connected to the battery by changing the switch position to B. Applying a battery voltage in this manner is equivalent to providing the circuit with a step-input excitation. We are particularly interested to know how the circuits behave, the instant the switch is changed to position B, and with the passage of time from that instant. We are also interested to know how the circuits behave when the switch is returned to position A, after it has been left in position B for an adequate length of time determined by considerations to be discussed later. While it is possible to investigate the manner in which both voltage and current associated with a given component change with time for any of the given circuits, we shall consider in detail only the current. The current waveforms for each circuit are shown in Figure 7.2.

Resistive Circuit

The diagram in Figure 7.1a shows a circuit which consists of a resistor, switch and a battery, connected in series. It must be emphasized at this stage that the components we are dealing with here are assumed to be ideal. An ideal component is one which only exhibits the property attributed to it. In practice the battery shown in our circuit will have an internal resistance, and the resistor will have some inductance, particularly if it is of a wire-wound type. Some stray capacitance will also be associated with it, and its effect will be significant in high

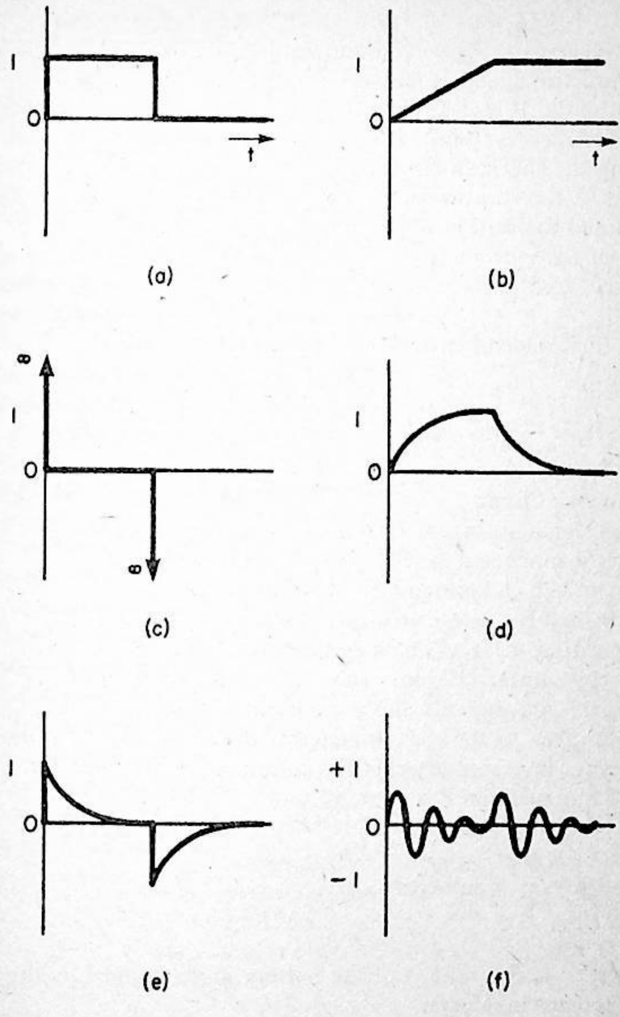


Fig. 7.2 Current waveforms for DC circuits in figure 7.1.

frequency applications. Investigating the behaviour of the circuits shown from an ideal component point of view helps us to appreciate what is going on, better.

The diagram in Figure 7.2a shows the current waveform for the resistive circuit from the instant the switch is moved to position B. It is clear from the diagram that the final value of current is established no sooner the switch position is changed. This is an important property of a purely resistive circuit. Let us now see what happens when the switch is returned to position A. From our earlier comments we would expect the current to drop to zero no sooner the switch is moved, and this is in fact what happens.

The final value of current with the switch in position B is given by

$$I = \frac{V}{R} \text{ amperes} \quad \ldots\ldots\ldots\ldots \quad 7.1$$

Inductive Circuit

The diagram in Figure 7.1b shows an inductor, a switch and a battery connected in series. The current waveform is shown in Figure 7.2b and gives a history of the current from the instant the switch is changed to position B and its subsequent return to position A. It will be seen that the current increases from zero in a linear fashion. There is nothing to oppose this increase and theoretically the current can go on increasing to infinity. As the current increases the energy stored in the inductor increases as well. The current at any instant with the switch in position B is given by

$$I = \frac{V}{L} t \text{ amperes} \quad \ldots\ldots\ldots\ldots \quad 7.2$$

where V is the voltage of the battery in volts and L is the inductance in henrys.

The energy stored in the inductor is given by

$$E = \tfrac{1}{2} L I^2 \text{ joules} \quad \ldots\ldots\ldots\ldots \quad 7.3$$

When the switch is returned to position A the current continues to flow in the same direction along the short circuit path available, and remains constant. The current is maintained by the energy stored in the magnetic field. When current flows in a closed path having no resistance, energy is not dissipated. Hence the energy stored in the magnetic field remains unaltered.

Capacitive Circuit

The diagram in Figure 7.1c shows a capacitor, a switch and a battery connected in a series fashion. The current wave form is shown in Figure 7.2c, and gives a history of the current from the instant the switch is moved to position B and subsequently returned to position A. Immediately the switch is moved to A, the current shoots up to infinity and returns to zero, in zero time. The name for a function behaving like this is an impulse. When the switch is returned to position B, a second impulse is generated, however on this occasion it shoots to infinity in the opposite direction. For the events just described what happens is that the capacitor charges up to the battery voltage in zero time, when the switch is moved to B. Once this has occurred the battery can no longer drive any current through the circuit since its voltage is exactly cancelled out by the voltage on the capacitor. Therefore the current is immediately reduced to zero. Returning the switch to A causes the capacitor to discharge, and this occurs in zero time. Naturally, once the capacitor has discharged completely, no current can flow in the circuit.

The voltage on the capacitor rises to the battery voltage no sooner the switch is moved to position B. It remains there until the switch is returned to position A, when it returns immediately to zero. The voltage on the capacitor takes the form of a step function.

Resistive-Inductive (R–L) Series Circuit

The diagram in Figure 7.1d shows a resistor, inductor, switch and a battery, connected in the form of a series circuit. The current waveforms are shown in Figure 7.2d and these are obtained when the switch is moved to position B and subsequently returned to position A. From this figure it is seen that the current starts to increase from zero at a finite rate, no sooner the switch is moved to position B. The rate of increase decreases with the passing of time, and it can be shown theoretically that it reduces to zero at infinite time. An important property of an R–L circuit is its time-constant, T. This is defined as the time taken for the current to reach its final value had the initial rate of increase been maintained. It also works out that the current reaches 63% of its final value in 1 time-constant. The time-constant T is given by

$$T = \frac{L}{R} \text{ seconds} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.4$$

where L is the inductance in henrys and R is the resistance in ohms.

This expression shows that the time-constant can be altered by changing the inductance, resistance, or both. A particular note should be made of the fact that increasing the resistance decreases the time-constant, and vice versa. Had the resistance been zero then the time-constant would have been infinity, and the circuit reduces to that of Figure 7.1b. Had the resistance been infinity then the time-constant would have been zero, and the circuit reduces to that of Figure 7.1a.

The current at any time t after the switch has been moved to position B is given by

$$i_t = \frac{V}{R}\left(1 - e^{(-t/T)}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.5$$

where V is the battery voltage in volts, R is the resistance in ohms, L is the inductance in henrys, T is the time constant in seconds and equal to L/R and e is equal to 2.718, a constant.

In equation 7.5 the exponential term $e^{(-t/T)}$ gets smaller and smaller as time passes, and eventually becomes zero after infinite time has elapsed. In practice the term is small enough to be neglected once a time equal to 4 time-constants has passed. The current is then limited only by the resistance present in the circuit and is given by

$$I = \frac{V}{R} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad 7.6$$

The voltage across the inductor at any time after the switch has been moved to position B is given by

$$V_L = V_B \, e^{(-t/T)} \quad \dots\dots\dots\dots\dots\dots\dots \quad 7.7$$

From this expression we note that V_L is a maximum and equal to the battery voltage at the instant the switch is moved to position B. This is so since putting $t = 0$ in the equation reduces the exponential part of the expression to unity, leaving $V_L = V_B$. Putting $t = \infty$ reduces the exponential part of the expression to zero, and describes the state of affairs after infinite time has passed. The voltage across the inductor is then zero.

Returning the switch to position A causes the current to continue flowing in the same direction as before. However unlike for the case with inductance by itself, the current now decays at a rate governed by the time-constant. It eventually reaches zero after infinite time. In practice the current will have reached approximately zero in 4 time-constants. Whenever an exponential growth or decay is involved this property may be assumed. The current at any time t after the switch has been returned to position A is given by

$$i_t \simeq \frac{V}{R} \, e^{(-t/T)} \quad \dots\dots\dots\dots\dots\dots\dots \quad 7.8$$

Putting t = 0 in this expression reduces the exponential part to unity and the current is simply given by the ratio of battery voltage to resistance. This is the current which flows in the circuit at the instant the switch is returned to position A. It is also the final value of current with the switch in position B.

The voltage across the inductor assumes a negative value no sooner the switch is returned to position A. It then decays exponentially to zero. The voltage at any time t after the switch has been returned to position A is given by

$$V_L = -V_B \, e^{(-t/T)} \quad \ldots\ldots\ldots\ldots \quad 7.9$$

Resistive-Capacitive (RC) Series Circuit

The diagram in Figure 7.1e shows a resistor, capacitor, a switch and a battery connected in a series fashion. The waveforms are shown in Figure 7.2e, and these are obtained when the switch is moved to position B and subsequently returned to position A. No sooner the switch is moved to position A the current assumes a value determined by the resistance, and given by straightforward application of Ohm's Law. This behaviour suggests that the capacitor behaves like a short circuit at the instant of switch on. As time passes the current decays at a rate governed by the time-constant, and eventually reaches zero at infinite time. The fact that the current reduces to zero indicates that the capacitor behaves as an open circuit after sufficient time has passed. The time-constant of a RC circuit is given by

$$T = CR \text{ seconds}$$

where C is the capacitance in farads, and R is the resistance in ohms.

When the switch is returned to position A the capacitor again behaves as a short circuit at the instant of switching, and as before the current is limited by the resistance. The waveform diagram shows that the voltage across the capacitor just before returning the switch to position A is almost equal to the battery voltage, having approached it exponentially from zero taking a time greater than 4 time-constants to do so. The current at any time t after returning the switch to position A will therefore be given by

$$i_t = -\frac{V}{R} e^{(-t/T)} \quad \text{.......................} \quad 7.10$$

This expression shows a reversal in the direction of current. While the current approaches zero from a negative value, the voltage across the capacitor will decay exponentially to zero from a positive value. The rate of change of both variable parameters will be governed by the time-constant.

The voltage across the capacitor at any time t when the switch is in position B is given by

$$V_C = V_B \left(1 - e^{(-t/T)}\right) \quad \text{.................} \quad 7.11$$

The voltage across the capacitor at any time t after the switch is returned to position A is given by

$$V_C = V_B \, e^{(-t/T)} \quad \text{.....................} \quad 7.12$$

We sometimes meet situations where an additional resistor is connected across the capacitor. The presence of this resistor merely reduces the final value of voltage to which the capacitor charges up, and the associated time-constant. The two resistors behave as a potential divider where the battery voltage is concerned. In calculating the time-constant their parallel equivalent is taken into be the effective resistance.

Resistive-Inductive-Capacitive (RLC) Series Circuit

The diagram in Figure 7.1f shows a resistor, an inductor, a capacitor, a switch, and a battery connected in a series fashion. The waveforms are shown in Figure 7.2f and are those obtained when the switch is moved to position B and subsequently returned to position A. The experienced hobbyist will

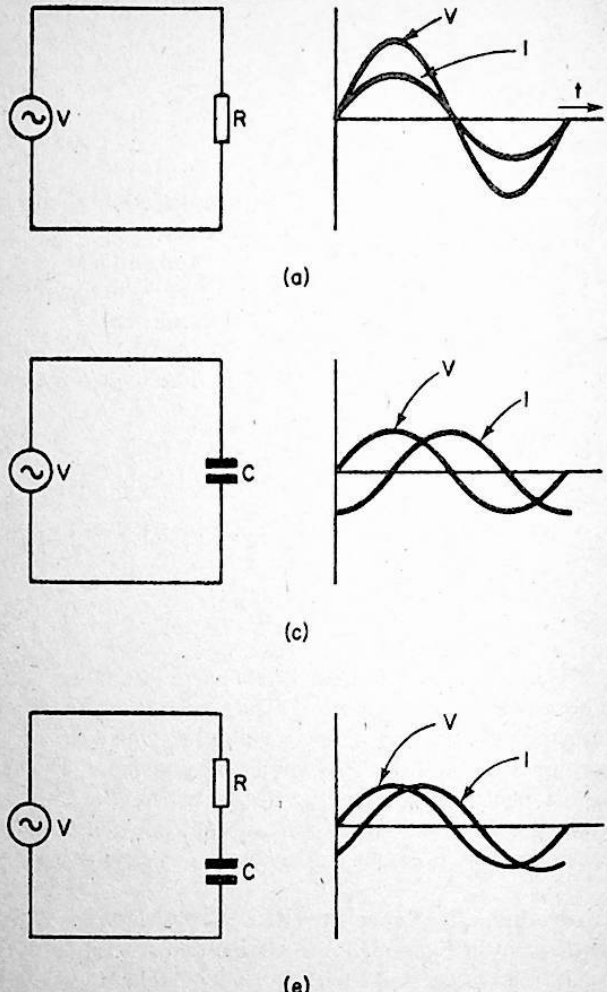
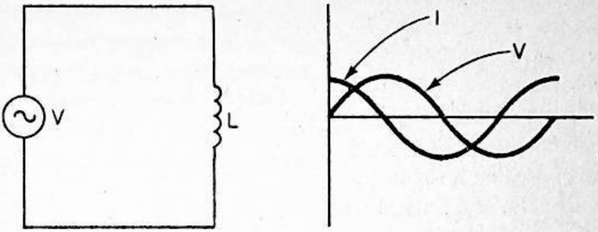
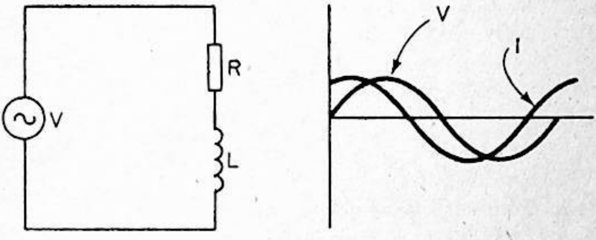


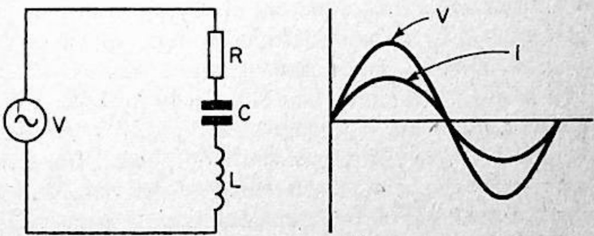
Fig. 7.3 AC circuits and waveforms.



(b)



(d)



At resonance

(f)

immediately identify this circuit as being the well known series resonant circuit. We are interested to know how this circuit behaves when it is excited by a step input voltage. On moving the switch to position A a sinusoidally varying current is established in the circuit. Due to the presence of resistance, the amplitude of each cycle will steadily decrease and eventually reach zero. The current will execute sinusoidal oscillations in this manner only if the resistance is less than a certain value known as the critical resistance. If the resistance is equal to the critical resistance the amplitude will immediately assume a maximum value and decay exponentially. If the resistance is greater than the critical value, the response will be very sluggish. It will labour along to a maximum and decay in a similar fashion. The hobbyist who has access to an oscilloscope will find it very absoring to verify this behaviour. Incidentally the value of critical resistance is given by

$$R_c = 2\sqrt{(L/C)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.13$$

AC Circuits

An ac circuit is one which is energized or excited by an alternating voltage. The preferred waveform for such a voltage is the sinusoid. The way in which circuits behave when subjected to a sinusoidal excitation, make the sinusoid ideally suited as a test waveform. A principal feature is that the voltages and currents present at any point in a linear circuit excited by a sinusoidal source, also happens to be sinusoidal. The excitation sinusoid and a sinusoid observed at any point in the circuit can differ only in respect of amplitude and phase. The diagram in Figure 7.3 helps us to understand the significance of the term phase. The term phase indicates to us the displacement of one waveform with respect to the other on the horizontal axis. It is measured in degrees or radians. Conversion from one to the other can be carried out by remembering that 2π radians is equal to 360°. π is a constant and is equal to 3.142. The phase difference may also be expressed in units of time, if the frequency is known. The mathematician, Fourier, has shown that any waveform irrespective of its complexity, consists of sinusoids suitably weighted in amplitude and phase. This is an extremely important fact. It immediately suggests that a linear circuit which is designed to handle a non-sinusoidal excitation, can be tested using a sinusoidal excitation.

The root-mean-square or rms value of a sinusoid is an important quantity which should receive our attention. Its significance stems from the fact that a sinusoidal voltage of maximum amplitude or peak value V_m, and a dc voltage equal to the rms value produces the same amount of heat energy when driving a current through a given value of resistance. The rms value of any sinusoid is obtained by dividing the maximum amplitude by 1.414. In equation form this may be stated as

$$V_{rms} = \frac{V_m}{1.414} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.14$$

If the rms value is known and we wish to find the maximum value it is simply necessary to multiply the rms value by 1.414. In equation form this may be stated as

$$V_m = 1.414\, V_{rms} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad 7.15$$

The best example of a sinusoidal source is the domestic mains supply. This supply has a rms voltage of 230 volts and alternates sinusoidally at a frequency of 50 Hz.

There are two aspects of ac circuit behaviour we must investigate with care. The first relates to the voltage and current relationship, and the second concerns the influence of frequency on the circuit current and on the voltage across circuit elements. The diagrams in Figure 7.4 show several circuits. Each of these circuits are excited by a sinusoidal source. The voltage and current waveforms applicable to each circuit are shown alongside.

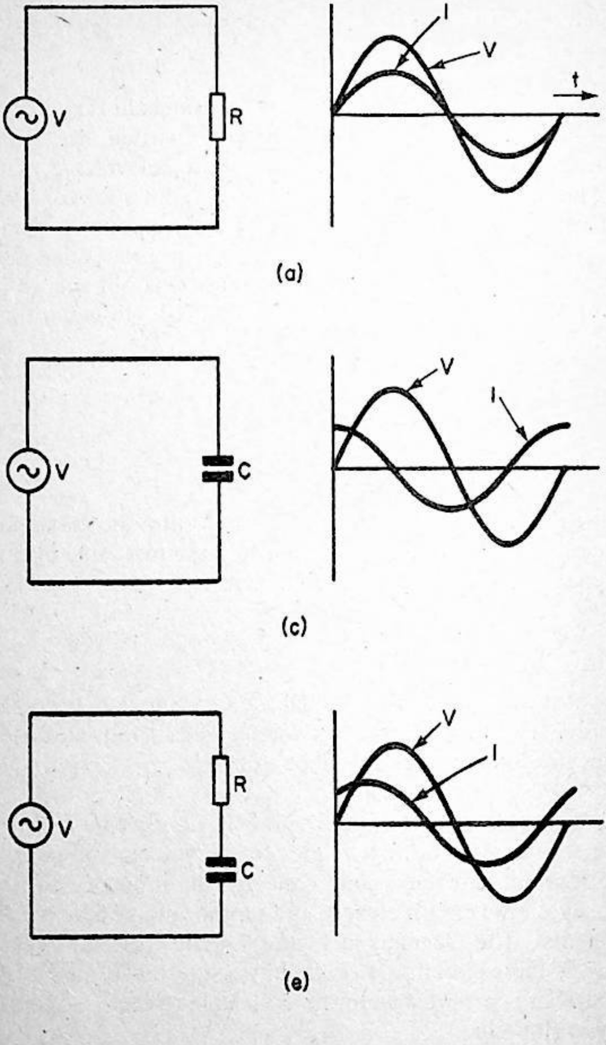
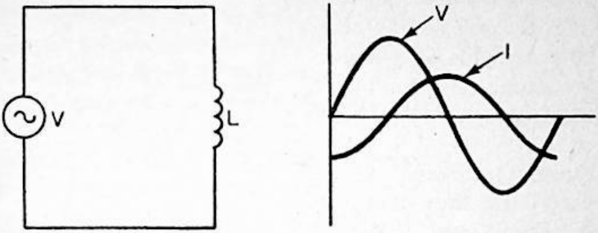
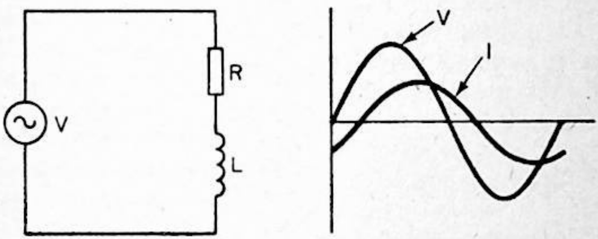


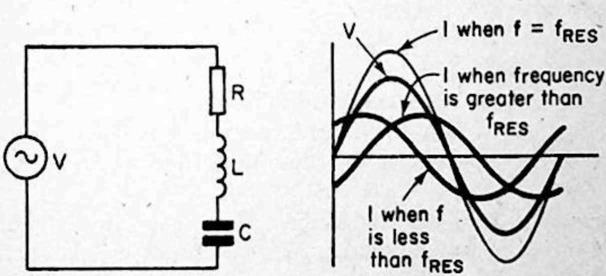
Fig. 7.4


V
L
V
I

(b)


R
V
L
V
I

(d)


R
V
L
C
V
I when $f = f_{RES}$
I when frequency is greater than f_{RES}
I when f is less than f_{RES}

(f)

The diagram in Figure 7.4a shows a purely resistive circuit. The resistor behaves very much like in a dc circuit. It functions to limit the current flow. A significant feature about this circuit is the so-called in-phase relationship between voltage and current. This is highlighted by the waveforms. The frequency has no effect on the current. The current is obtained by a straightforward application of Ohm's Law.

The diagram in Figure 7.4b shows a purely inductive circuit. The waveforms show that the voltage and current are not in phase. The current in fact lags the voltage by 90° or $\frac{\pi}{2}$ radians. This behaviour is fundamental to an inductor and should be remembered. The magnitude of the current is influenced by frequency. The opposition to the flow of current increases with frequency. Inspite of these observations we may still apply Ohm's Law to find the current. However now, the resistance R in the denominator is replaced by the inductive reactance. The inductive reactance X_L is given by

$$X_L = \omega L \quad \ldots\ldots\ldots\ldots \quad 7.16$$

Substituting this value for R in the Ohm's Law expression gives

$$I = \frac{V}{\omega L}$$

This expression for current shows that the current is inversely proportional to frequency and inductance. An increase in either of these quantities will cause the current to decrease.

The diagram in Figure 7.4c shows a purely capacitive circuit excited by a sinusoidal source. The waveforms indicate that the voltage and current are again not in phase. However in this case the current leads the voltage by 90° or $\pi/2$ radians. The magnitude of the current is governed by the value of capacitance and frequency. Ohm's Law may be used as before to obtain the current. All we need to do is to replace the resistance R in the familiar expression by the capacitive reactance X_c. The capacitive reactance is given by

$$X_c = \frac{1}{\omega C} \quad \text{7.17}$$

Substituting for X_c in the Ohm's Law expression gives

$$I = \omega CV \quad \text{7.18}$$

This expression shows that the current is directly proportional to frequency and capacitance. An increase in either would cause the current to increase.

So far we have only considered the behaviour of each passive element on its own. Several interesting points emerge when we combine these elements. Let us first consider the RL combination. The diagram in Figure 7.4d shows a RL circuit excited by a sinusoidal source. From the waveforms we note that the voltage and current are no longer exactly 90° out of step. They are out of phase by an amount less than this figure. The presence of resistance in other words has forced the phase difference to be less than 90°. An interesting situation occurs when the resistance is numerically equal to the reactance. The phase difference now becomes 45°. Further increase in resistance causes this angle to reduce still further, and finally when the resistance is extremely large in comparison to the reactance, the angle approaches zero. The circuit would then behave as if though it was a pure resistance. In order to calculate the current in a RL circuit we must first determine its impedance. The magnitude of the impedance is given by

$$|Z| = \sqrt{(R^2 + \omega^2 L^2)} \quad \text{7.19}$$

The current may be obtained by substituting $|Z|$ for R in the Ohm's Law expression. Thus we have

$$I = \frac{1}{\sqrt{(R^2 + \omega^2 L^2)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.20$$

The diagram in Figure 7.4e shows a RC circuit, excited by a sinusoidal source. From the waveforms we note that the voltage and current are no longer exactly 90° out of phase. While the current still leads the voltage it is out of step by an angle which is less than 90°. When the resistance is equal to the capacitive reactance the phase difference reduces to 45°. A further increase of resistance causes the phase difference to decrease still further. The phase difference approaches zero when the resistance is very large in comparison to the capacitive reactance. The circuit then behaves as if though it was a pure resistance. To determine the current we need to know the magnitude of the impedance. The magnitude of impedance for a RC circuit is given by

$$|Z| = \sqrt{\left(R^2 + \frac{1}{\omega^2 C^2}\right)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.21$$

Replacing R by | Z | in the well-known Ohm's Law expression gives

$$I = \frac{V}{\sqrt{\left(R^2 + \frac{1}{\omega^2 C^2}\right)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 7.22$$

The diagram in Figure 7.4f shows a RLC circuit excited by a sinusoidal source. This is the well known series resonant circuit. This circuit has the remarkable property of behaving as a pure resistance at a given frequency. At this frequency the inductive and capacitive reactances cancel each other leaving only the resistance. As we would expect the supply voltage and current are exactly in phase at this frequency. This frequency is called the resonant frequency. If the frequency is increased above the resonant frequency, the circuit inductance predominates and the current lags the supply voltage. If the frequency is decreased below the resonant frequency, the circuit capacitance predominates and the current leads the voltage. The importance of this circuit is due to its ability to select a given frequency. The circuit demands the maximum current from the supply at resonance, that is when the inductive reactance cancels out the capacitive reactance. By decreasing the resistance we could improve the ability of the circuit to select a given frequency and reject sharply frequencies on either side of it. The magnitude of the impedance at any frequency is given by

$$|Z| = \sqrt{\left\{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2\right\}} \quad \ldots\ldots\ldots\ldots \quad 7.23$$

The current at any frequency may be found by using this expression in the Ohm's Law equation. Thus

$$I = \frac{V}{\sqrt{\left\{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2\right\}}} \quad \ldots\ldots\ldots\ldots \quad 7.24$$

The current at resonance is given simply by

$$I_R = \frac{V}{R} \quad \ldots\ldots\ldots\ldots \quad 7.25$$

Illustrative Examples

Example 7.1

A 10 kΩ and a 4.7. kΩ resistor are connected across a 10 V dc supply, in a series fashion. Calculate the circuit current and the voltage across each resistor.

Solution

The circuit current is obtained by application of Ohm's Law. Therefore

$$I = \frac{V}{R} = \frac{10}{10K + 4.7K} = 0.68 \text{ mA}$$

The voltage across each resistor is given by the product of current and resistance, and involves again an application of Ohm's Law.

Thus we have

$$V = IR$$

The voltage across the 10 k resistor is therefore given by

$$V = 0.68 \text{ mA} \times 10 \text{ k} = 6.8 \text{ V}$$

and the voltage across the 4.7 k resistor is given by

$$V = 0.68 \text{ mA} \times 4.7 \text{ k} = 3.2 \text{ V}$$

Example 7.2
A series RC circuit consists of 1 MΩ resistor and a fully discharged 1 μF capacitor. Find the time constant of the network and calculate the initial current which flows when the network is connected to a 10 V supply.

Solution
The time constant for a series RC circuit is given by

$$T = CR$$

Substituting values

$$T = 1.0 \times 10^{-6} \times 1.0 \times 10^{6} = 1 \text{ sec}$$

The initial current is given by

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

Substituting the values given

$$I = \frac{10}{1 \times 10^6} \text{ amps} = \frac{10}{1 \times 10^6} \times 10^6 \ \mu A$$

$$= 10 \, \mu A$$

Example 7.3

An alternating voltage has a peak value of 6 V. Find its rms value.

Using equation 7.14,

$$V_{rms} = \frac{V_p}{1.414}$$

Therefore inserting values given for V_p

$$V_{rms} = \frac{6}{1.414} = 4.24 \text{ V}$$

Example 7.4

A series circuit consists of a 10 ohm resistor, and a 1 mH inductor. Calculate the reactance of the inductor at a frequency of 1 kHz, and the circuit impedance at the same frequency.

Solution

From equation 7.16, the reactance is given by

$$X_L = 2\pi fL$$

Inserting values we have

$$X_L = 2 \times 3.142 \times 1 \times 10^2 \times 1 \times 10^{-3} = 6.28 \text{ ohms}$$

From equation 7.19, the magnitude of the impedance is given by

$$|Z| = \sqrt{(R^2 + \omega^2 L^2)}$$

Inserting values

$$Z = \sqrt{(10^2 + 6.28^2)} = 11.8 \text{ ohms}$$

Chapter 8

USING DIODES AND TRANSISTORS

Using Diodes

The diode can perform many useful circuit functions. In all these applications, the one feature which is exploited, is the diode's ability to conduct electricity in one direction and not in the other. Rectification, clamping, clipping, signal routing, and source/load isolation are a few applications where the diode is prominently encountered.

In using a junction diode there are a few important considerations that require our attention. A diode in its basic form consists of a pn junction. A depletion layer is always present on either side of a pn junction, and in biasing the device we merely modify the width of this layer. When the junction is forward biased the width is reduced and when the device is reverse biased the width is increased. The manufacturer always specifies the maximum current that the device would safely handle in the forward biased condition, the maximum reverse voltage that could be applied in the reverse biased condition and the power dissipation. From our earlier work we know that exceeding the reverse voltage beyond a certain value can cause the junction to breakdown. In voltage reference diodes this behaviour works to our advantage while in a general purpose diode it is a distinct disadvantage. Ideally we would like a general purpose diode to behave as a short circuit in the forward biased condition and an open circuit in the reverse biased condition. In the forward biased condition it should handle any current we please to pass through it and in the reverse biased condition it should safely accept any voltage we care to apply across it. In practice, the maximum current is limited by the maximum power dissipation, and the maximum reverse voltage is limited by the junction breakdown voltage. A typical switching diode will be able to handle a mean forward current of 75 mA, and withstand a reverse voltage of the same order. Diodes used in rectifier applications will naturally be able to handle larger currents and also withstand higher maximum reverse voltages. In rectifier applications using capacitive smoothing, we need to watch the peak current rating and the peak inverse voltage rating. At switch on, the capacitor behaves as a short circuit and can demand a high current for a brief moment. The diode should be able to handle this current safely. During the non-conducting period, the voltage across the diode can be as much as twice the peak value of the alternating input voltage. The maximum reverse voltage rating, or the peak inverse voltage rating needs to be higher than twice the peak value of the input alternating voltage.

In switching applications we need to ensure that the voltage waveforms presented to the diode have the desired effect. If it is our intention to forward bias a diode when a given voltage assumes a positive value, the diode will have to be connected as shown in Figure 8.1. In this circuit the diode will become forward biased no sooner the input voltage rises above 0.6 V. When the input voltage assumes a negative value the diode will become reverse biased. By reversing the terminals of the diode we have the situation where the diode

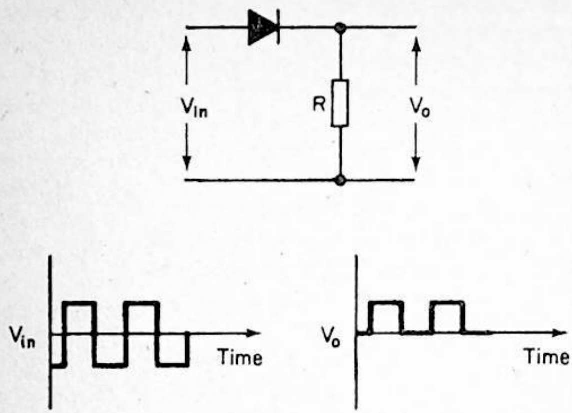


Fig. 8.1 Using the diode as a switching element.

conducts when the input assumes a negative value, and cuts-off when the input assumes a positive value. By returning the diode to some voltage other than zero, we are able to define the voltage at which it commences to conduct.

Using Transistors

The first step in using a transistor is to decide on suitable biasing scheme. The choice will depend heavily on the application. If the requirement is for an amplifier the biasing scheme should ensure that the device operates in a linear mode. If the requirement is for a switch then the biasing arrangement must ensure that the device is either on or off. In other words it will now be functioning in a non-linear mode. We shall discuss first the possibilities open to us when we wish to use the device in a linear mode.

The simplest method of biasing is shown in Figure 8.2. The base-emitter junction is forward biased by returning the base to a voltage positive with respect to the emitter. Had the device been of the pnp type then the base and collector

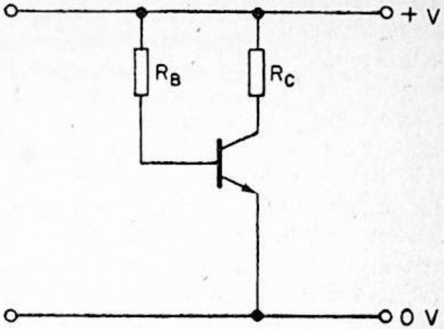


Fig. 8.2 Simple biasing scheme.

would have been returned to a negative voltage with respect to the emitter. Now the voltage across a forward biased junction is typically 0.6 V for silicon and 0.3 V for germanium. Hence the base voltage must necessarily be negative with respect to the collector, for the npn device shown. Therefore the base to collector junction is reverse biased. The collector current is a function of the base current. Operation in a linear fashion implies that a change in base current in either direction about a given value causes a proportionate change in the collector current. The resistance in the base circuit is dimensioned to satisfy this condition. An output voltage proportional to the collector current is obtained by employing a resistor in the collector circuit. This resistor has to be dimensioned such that only part of the available supply voltage is dropped across it. Only by satisfying this condition are we able to make sure that the output voltage bears a linear relationship to the input voltage. This simple circuit while giving satisfactory performance with silicon transistors is not suitable for use with germanium devices. The leakage current in a germanium transistor is much higher than that of a silicon transistor and doubles for every 10°C increase in temperature. As a consequence the collector current is heavily dependent on temperature, and can easily cause the operating point to drift with temperature. The operating point is specified by specifying a collector-emitter voltage and a collector current. The simple circuit we have just considered, is not capable of minimizing the influence of leakage current on the operating point, and due to this reason it is not recommended for use with germanium transistors. The leakage current does not pose a problem with silicon transistors. While it increases at a faster rate than that of germanium, it is initially so small that we could almost ignore its presence. However the base-emitter voltage of a silicon transistor is rather temperature sensitive, and this fact has to be recognized when deciding on a biasing scheme. Any scheme, where changes in base to emitter voltage does not introduce a change in collector current, is acceptable. From this viewpoint our simple circuit is acceptable.

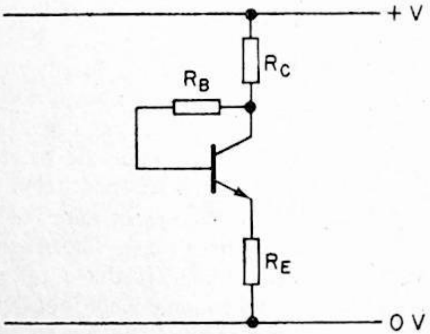


Fig. 8.3 Biasing scheme employing current and voltage negative feedback.

By adding one resistor and returning the base resistor to the collector terminal, a biasing scheme with improved stability is obtained. The circuit diagram is shown in Figure 8.3. This circuit arrangement employs both voltage and current negative feedback. Feedback is the name given to the process whereby changes taking place at a given point in a circuit or system is fedback to an earlier point in such a manner as to modify the overall behaviour. If the feedback acts to accentuate the original change, it is said to be positive feedback. If the feedback acts to arrest the original change, it is said to be negative feedback. By returning the base resistor to the collector, we monitor changes in collector voltage, and in doing so force the base current to adjust itself in such a manner as to correct for the original change in the collector voltage. This is an example of shunt voltage negative feedback. By adding a resistor in series with the emitter we are able to incorporate current feedback as well. Again the feedback is negative. The current which flows in this resistor causes a voltage to develop across it, and this voltage acts to control the effective base-emitter voltage. If this current should change for some reason, say a change in temperature, the negative feedback process will act to immediately correct that change, by modifying the base-emitter voltage. The emitter current can be controlled in this manner since it is a function of the base-emitter voltage. In this process we are effectively controlling the collector current since it is almost equal to the emitter current. We now appreciate why this circuit is superior to the simple circuit

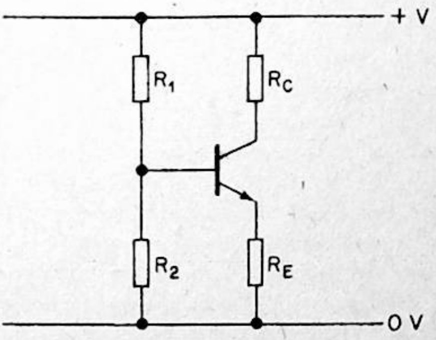


Fig. 8.4 Superior biasing scheme employing current feedback.

discussed earlier and also understand its operation from a negative feedback point of view.

A biasing scheme which requires yet another resistor, and whose performance is acceptable for most applications, is shown in Figure 8.4. This circuit employs only current

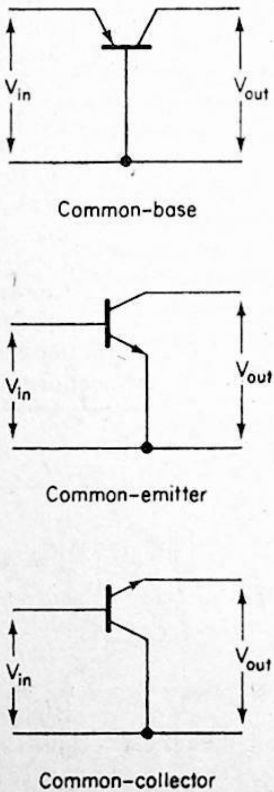


Fig. 8.5 The three operating modes for employing a transistor to handle signals.

negative feedback. Its superior performance is a result of the potential divider network in the base circuit. The presence of this network makes the collector current less dependent on the leakage current and hence this biasing scheme is particularly suited for use with germanium transistors.

Having considered the important biasing schemes, let us now see how the transistor can be employed as an amplifying element. There are three ways in which the transistor can be connected to handle signals. These are identified by the terminal which is common to both input and output. Thus we have three modes, namely, common-base, common-emitter, and common-collector to consider. These modes are illustrated in Figure 8.5. The bias circuits are excluded since we are now investigating the behaviour only from a signals viewpoint. The common-base arrangement is capable of giving high voltage amplification. The input resistance is rather low and this is a significant disadvantage for certain applications. The common-emitter arrangement is by far the most popular, since it provides both voltage and current amplification. Its input resistance can be conveniently arranged to meet a given requirement by application of negative feedback. The common-collector arrangement better known as the emitter-follower, is of special value as an impedance transforming stage. It is characterized by a high input resistance and low output resistance.

The diagram in Figure 8.6 shows a common-emitter amplifying stage. The biasing scheme used is that shown in Figure 8.3. We notice that the base resistor has been split into two parts and the mid-point returned to the signal common point. This has been done to prevent signal negative feedback from taking place. A feature of shunt negative feedback is to reduce the input impedance and this will be undesirable for an amplifier that should operate down to very low frequencies. A low input resistance will require the input capacitor to be excessively large.

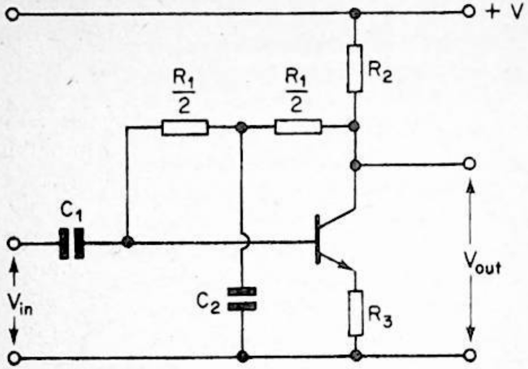


Fig. 8.6 A typical transistor amplifying circuit.

The low-frequency cut-off of our stage is given by

$$f = \frac{1}{2\pi R_{in} C_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 8.1$$

The input resistance R_{in} can be estimated in the following way. This resistance consists of two parts in parallel. The first part consists of the resistance seen looking into the base terminal with the bias resistor being ignored. This is essentially given by the product of the common-emitter current amplification factor, h_{fe}, and the total undecoupled resistance in the emitter leg. This resistance consists of r_e, the intrinsic emitter resistance, and R_3.

Since I_e is known, r_e can be evaluated using an expression introduced in Chapter 5, equation 5.22. If R_3 happens to be very much larger than r_e, then r_e could be ignored. The second part making up the input resistance simply is $R_1/2$. This of course assumes that the decoupling capacitor C_2 has a reactance which is very low compared to $R_1/2$ at the lowest frequency of interest. These two parts in parallel comprise the input resistance. Equation 8.1 can be rearranged to give C in terms of the input resistance and frequency, and can be used to calculate the required value of capacitance to obtain a desired 3 dB frequency. The rearranged equation is

$$C = \frac{1}{2\pi f R_{in}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 8.2$$

So far we have discussed the aspects which conern us when using the transistor in a linear role. It is possible to use the transistor in a non-linear role as well. The device is then said to function as a switch. The circuit diagram in Figure 8.7

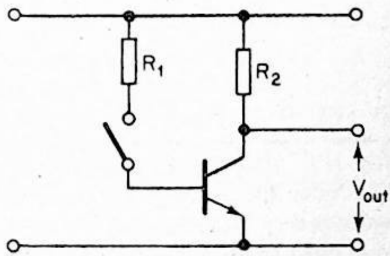


Fig. 8.7 A typical transistor switching circuit.

shows how a transistor can be connected to function as a switch. When the switch in the base circuit is open and no current flows into the base terminal, the transistor is cut-off, and therefore no collector current flows. The voltage at the collector terminal is almost at the supply voltage. When the switch is closed, the base current which flows, causes the transistor to come on. The transistor is said to be on, when the voltage across the collector and emitter is low and equal to the saturation voltage of the device. In this condition almost all the supply voltage is dropped across the resistor R_L.

So far we have considered the problems associated in using bipolar transistors. Perhaps we should consider the essential points relating to unipolar transistors as well. Since the principle of operation is significantly different to the bipolar device, the biasing scheme differs. The diagram in Figure 8.8

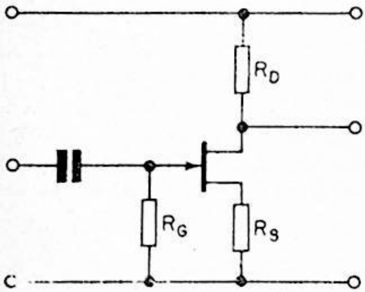


Fig. 8.8 Typical common–source JFET amplifying stage.

shows a n-channel JFET, biased for operation as a common source amplifier. Since the input junction is reverse biased the input resistance is very high, and this feature is recognized and appreciated in most linear unipolar transistor applications. The degree of reverse bias is governed by the value of the resistance in the source circuit. A given source current will produce a given voltage across this resistor, and this will be the effective voltage between the gate and source. Since the current which flows in the gate circuit is minute and is entirely due to leakage effects, the resistance between the gate and the zero volt rail can be dimensioned to be considerably high, without interfering with the chosen operating point. Thus the capacitor required at the input, for a given low frequency response, can be made considerably small.

When using any transistor, it is of utmost importance to make sure that the manufacturer's ratings are not exceeded. Particular attention should be directed to ensure that the operating drain current and the operating voltages present across electrodes are within the rated values. We should be equally well concerned about the power rating of the device. When using power transistors it is vital that an efficient heat sink is used to rapidly transfer the heat to the surrounding air. The manufacturer's rating should always be carefully studied, making a particular note of the conditions for which it is valid.

Illustrative Examples

Example 8.1

The cathode of a silicon junction diode is connected to the negative terminal of a 10 V battery. Its anode is returned via a 8.2 k resistor to the positive terminal of the battery. Determine the diode current.

Solution

Since the cathode is returned to the negative terminal and the anode is returned to the positive terminal, the diode is forward biased. The voltage drop across a forward biased silicon diode is 0.6 V. The current is limited by the series resistance, and is obtained by dividing the voltage across the resistor by the resistance.

Therefore

$$I = \frac{10 - 0.6}{8.2}\ \text{mA}$$

$$= \frac{9.4}{8.2}$$

$$= 1.15\ \text{mA}$$

Example 8.2
It is required to bias a transistor so that its collector-emitter voltage is 3 V, and the collector current is 1 mA. The biasing scheme should use only one resistor in the base circuit, and the circuit must operate from a 10 V supply. The transistor has a common-emitter current amplification factor of 100 at 1 mA collector current.

Solution
The required biasing scheme is shown in Figure 8.2 of the text. The base current must be found first. This may be obtained by using equation 5.20.

$$\therefore \quad I_b = \frac{I_c}{h_{fe}} = \frac{1}{100}\ \text{mA} = 10^{-2}\ \text{mA}$$

The voltage drop across the base-emitter junction is 0.6 V. Therefore the voltage available to force the base current through the base resistor is 10 V – 0.6 V. Using Ohm's Law

$$R_B = \frac{10 - 0.6}{10^{-2}\ \text{mA}}\ \text{k}\Omega = \frac{9.4}{10^{-2}}\ \text{k}\Omega = 940\ \text{K}$$

The nearest preferred value is 1 MΩ.

Since the collector-emitter voltage has to be 3 V, 10 V – 3 V, i.e. 7 V, has to be dropped across the collector resistor. Using Ohm's Law again,

$$R_C = \frac{7}{1\ \text{mA}} = 7\ \text{K}\Omega$$

The nearest preferred value is 6.8 KΩ.

Example 8.3
An amplifying stage has an input resistance of 47 k. The lower 3 dB frequency is set up by dimensioning the capacitor at the input. What value of capacitance would give a 3 dB point at 40 Hz?

Using equation 8.2, the capacitance is given by

$$C = \frac{1}{2\pi f R_{in}} = \frac{1}{2\pi \times 40 \times 47 \times 10^3} \text{ F}$$

$$= \frac{10^2 \ \mu F}{2\pi \times 40 \times 47 \times 10^3} = 0.085 \ \mu F$$

Example 8.4
A common-emitter amplifying stage employs an undecoupled resistor in the emitter circuit. The transistor h_{fe} is 100, at the operating point, current. Find the input resistance of the stage, ignoring the shunting effects due to the biasing scheme, employed. Comment on the effect of emitter resistor decoupling.

Solution
A resistor is said to be undecoupled when signal voltages can develop across it. By connecting a capacitor, which has a resistance, at least a tenth of the value of the resistance at the lowest frequency of interest, decoupling is achieved.

Now the presence of undecoupled resistance serves to increase the input resistance. Using equation 5.21

$$R_{in} = h_{fe} \ (R + r_e)$$

The intrinsic emitter resistance, r_e is a function of the collector current, and is arrived at using equation

$$r_e = \frac{25}{1 \text{ mA}} = 25 \ \Omega$$

and

$$R_{in} = 100(25 + 100) = 12500 = 12.5\ k\Omega$$

Decoupling the emitter resistance has the effect of reducing the input resistance, and increasing the stage gain. The gain is given by the ratio of the collector load resistance to the undecoupled emitter resistance, provided the collector resistance is not too large compared to the emitter resistance.

Chapter 9

USING INTEGRATED CIRCUITS

Linear Integrated Circuits

Integrated circuits like discrete devices need to be provided with suitable dc power supplies. The manufacturer specifies the operating voltage for his device and this should never be exceeded. In using integrated circuits an understanding of the conditions which exist at the terminals of the device, is invaluable in exploiting to the full its capabilities. This would assist in the correct selection of external components, for use with the chip. Due to its wide popularity and almost limitless list of applications, we shall concentrate on the operational amplifier. The operational amplifier or op-amp as it is better known is high gain amplifier capable of performing mathematical operations. Had this been its only virtue it would not be so popular. The features that make it suitable as an op-amp, also make it suited for a host of other applications. These applications include both linear and non-linear circuit functions. It is clearly not possible to discuss each and every application here. Therefore we shall limit ourselves to a few which the hobbyist is more likely to meet. Amplification of a signal is a common requirement, and the op-amp can be easily arranged to perform this function. We shall investigate how an op-amp is used for this purpose. Waveform generation is another common requirement, and here too an op-amp comes in useful. We shall have something to say on this aspect as well.

An op-amp has two input terminals and usually one output terminal. One input terminal is called the inverting input and the other is called the non-inverting input. These are indicated respectively by a – and + sign. This simply identifies the relationship of the output to the input, in terms of signal polarity. If a signal at the inverting input is positive going, then we would expect the signal at the output to be negative going. Similarly if the signal at the non-inverting was positive going we would expect the signal at the output to be positive going as well. In using the op-amp as an amplifier we could employ either input. If we use the non-inverting input the gain is given by

$$A = \frac{R_f}{R_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 9.1$$

The circuit diagram is shown in Figure 9.1. This expression for gain is very straightforward to use and should cause no problems at all. The usual procedure is to assume a convenient and suitable value for R_1 and calculate R_f for the required value of gain.

The value of R_2 should be approximately equal to the parallel equivalent of R_1 and R_f, if the input to the stage is from a low source resistance and is dc coupled. If the stage is ac coupled then it is right that we should make R_2 equal to R_f.

If we use the non-inverting input, the gain is given by

$$A = 1 + \frac{R_f}{R_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad 9.2$$

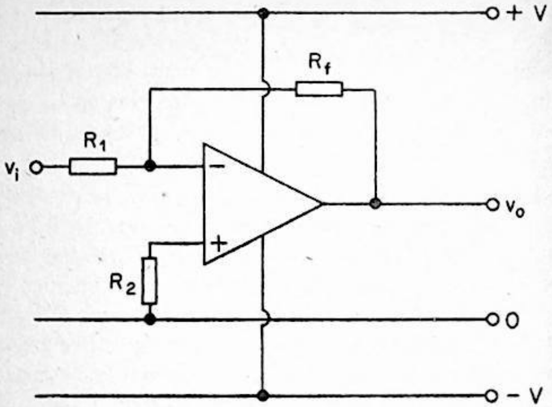


Fig. 9.1 Operational amplifier connected in an inverting mode.

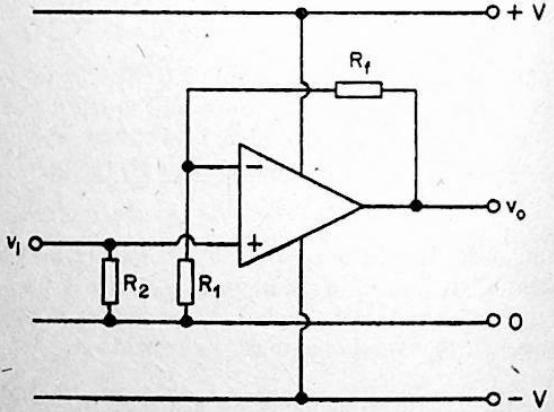


Fig. 9.2 Operational amplifier connected in a non-inverting mode.

The circuit diagram is shown in Figure 9.2. The expression is very similar to equation 9.1, with just an unity term added to the right-hand side. If the required stage gain is high then the resulting error from neglecting the unity term is small. The approach for calculating the values of resistance is as before.

We often have a requirement for an amplifier with a given frequency response. The op-amp lends itself well towards meeting this need. By a careful choice of RC networks, taking the place of R_f or R_1, or both, the desired response can often be achieved. Suppose we required an amplifier which had a response that increased at 6 dB per octave below a given frequency. This is easily obtained by the configuration shown in Figure 9.3. The break frequency or in other words the frequency at which the response would increase by 3 dB is given by

$$f = \frac{1}{2\pi C R_f} \qquad \text{9.3}$$

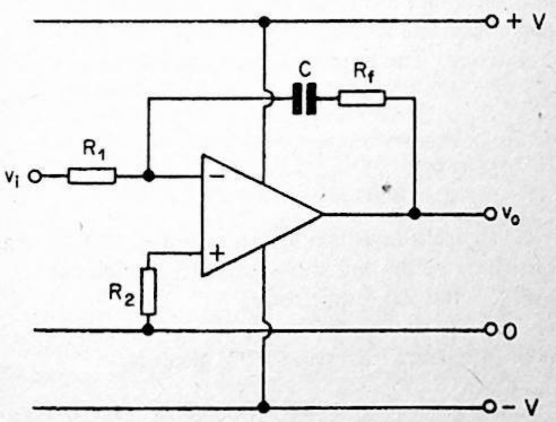


Fig. 9.3 Using an operational amplifier to give a rising response below a given frequency.

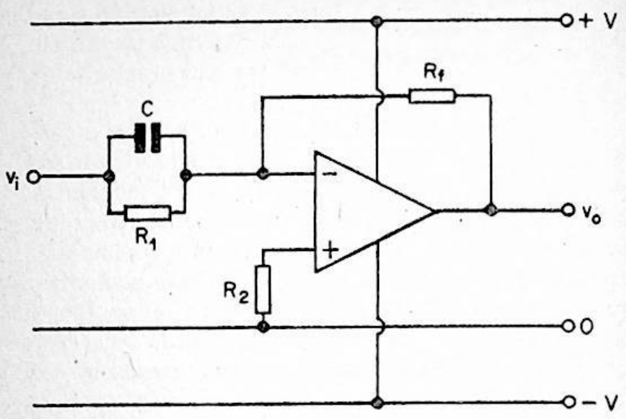


Fig. 9.4 Using an operational amplifier to give a response which rises above above above a given frequency.

The circuit diagram in Figure 9.4 shows the arrangement for achieving a response which increases with frequency above a given frequency. The break frequency in this case is given by

$$f = \frac{1}{2\pi C R_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots 9.4$$

Circuit configurations which give a response that decreases with frequency below and above given frequencies are shown in Figures 9.5 and 9.6 respectively.

The break frequency for Figure 9.5 is given by

$$f = \frac{1}{2\pi C R_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots 9.5$$

The break frequency for Figure 9.6 is given by

$$f = \frac{1}{2\pi C R_f} \qquad \text{9.6}$$

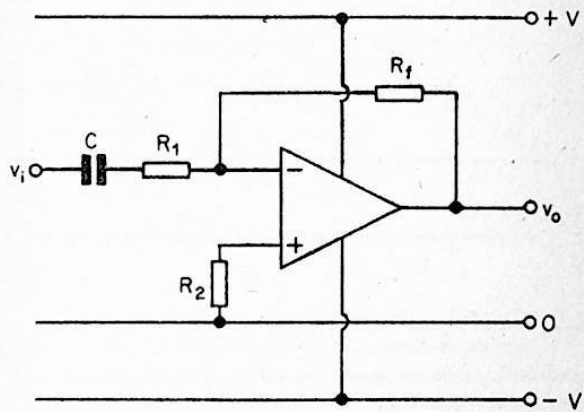


Fig. 9.5 Use of operational amplifier to give a response which decreases below a given frequency.

The use of the op-amp is greatly facilitated when positive and negative dc supplies are available. Sometimes these may not be available and we would like to use it from a single rail. This can be done by resorting to the arrangement shown in Figure 9.7. Capacitors C_1 and C_2 serve a dc blocking function. The value of C_1 should be such that its reactance is small in comparison to R_2 at the lowest frequency of interest. Likewise the reactance of C_2 should be small in comparison to R_1, at the lowest frequency of interest.

The operational amplifier may be employed for generating a variety of waveforms. The circuit diagram in Figure 9.8 shows how an op-amp is connected to generate sinewaves. Any

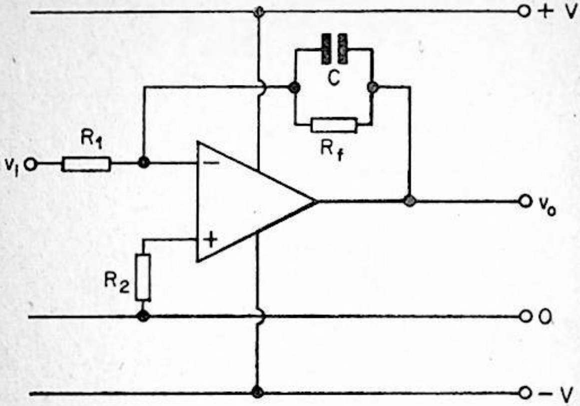


Fig. 9.6 Use of operational amplifier to give a response which decreases above a given frequency.

oscillator basically consists of an amplifier and a positive feed-back network. A proportion of the output voltage is fedback to the input and in doing so oscillations are produced and sustained, if the amplifier has a gain in excess of the minimum requirement. The feedback network in the circuit shown is the well known Wien network. Since the phase shift due to the network is zero at the oscillation frequency, the amplifier should produce no phase inversion. This condition is met by utilizing the non-inverting terminal. In an oscillator of this type, the frequency of oscillation is given by

$$f = \frac{1}{2\pi C R} \qquad \text{9.7}$$

The amplifier is required to have a minimum gain of 3. Utilizing the non-inverting input, the negative feedback network is dimensioned to satisfy this condition. In this manner, the amplitude may be stabilized to give a sinewave,

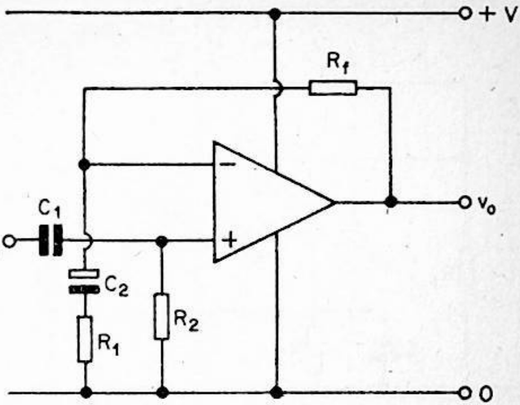


Fig. 9.7 Method of connecting operational amplifier to operate from single supply rail.

only at one frequency. If the Wien network component values are switched to produce many frequencies, the amplitude will have to be stabilized using a thermister in the negative feedback loop.

The operational amplifier may also be employed to produce square waveforms. Once again positive feedback is applied. However in this case the feedback network simply consists of a resistive network. The periodic time of the output waveform is determined by a RC network. The circuit arrangement is shown in Figure 9.9.

Digital Integrated Circuits

TTL

A variety of logic functions can be performed using chips from the TTL family. We shall consider those circuit functions which are likely to be of interest to the hobbyist. The logic

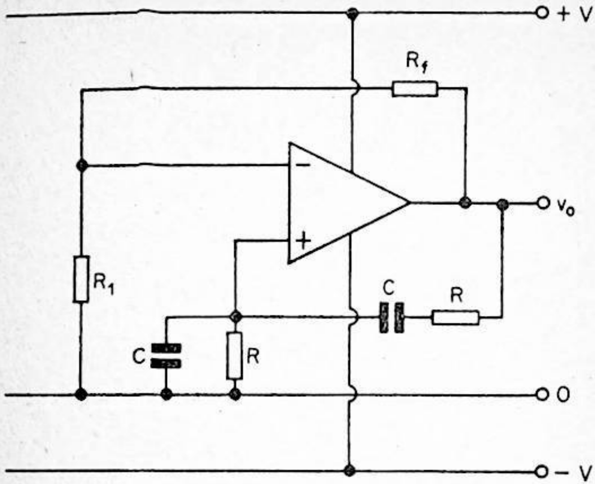


Fig. 9.8 Using an operational amplifier for sinusoidal waveform generation.

functions, OR, NOR, AND, NAND are performed using combinational logic gates. A gate is a device which has two or more inputs and one output, and behaves as a logic element. An AND gate yields an output when all the inputs are at a given logic level. Let us consider a two input AND gate. Suppose the inputs are identified by A and B. Now this gate will produce an output only and only if A AND B are at a specified logic value. From similar reasoning an OR gate will produce an output only and only if A OR B are at a specified logic level. The NOR and NAND functions are obtained when the outputs of the basic OR and NOR gates are inverted or negated. Thus a NAND gate simply consists of a AND gate followed by an invertor. Similarly an OR gate simply consists of an OR gate followed by an invertor.

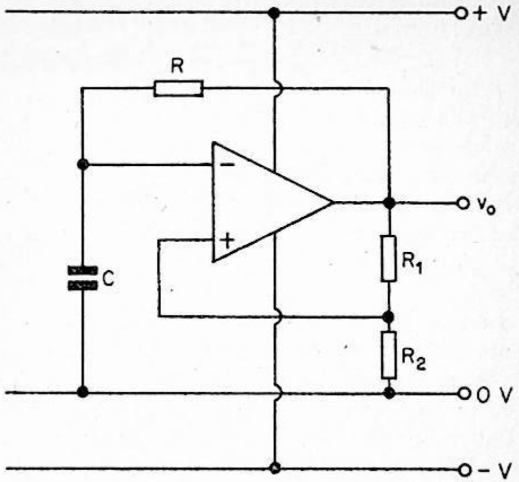


Fig. 9.9 Use of operational amplifier to generate square waves.

A logic chip normally comes in a plastic package, having fourteen terminals. The number of gates and the number of inputs that each gate can have is governed by the number of terminals available. Quad-2 input, Triple-3 input, Dual-4 input, and single-8 input gates are available in this type of package.

Depending on the circuit function, a sequential logic chip can come in a package having 14, 16, 24 or 40 terminals. A very large variety of sequential logic functions may be implemented using chips from the TTL family.

TTL operates from a nominal supply voltage of 5 V. The manufacturer requires that the supply voltage remains within ± 5% of 5 V, and care should be exercised to ensure that this condition is met. The power drawn from the supply depend on the nature of the logic circuit and can be of the order of tens of milliwatts. When using gates it is necessary to ensure that unused inputs are connected to one of the used inputs, or returned to the +5 V rail. An AND or OR gate may be used as an invertor by strapping the inputs together.

Members of a logic family can be connected to each other without upsetting the individual behaviour. When connecting together gates we must be aware of fan-out considerations. TTL has a fan-out of 10. The fan-out number refers to the number of inputs that can be connected to the output of a given device, without altering the logic levels.

CMOS

CMOS is characterized by low power consumption and its ability to operate from supply voltages which lie between 3 and 15 V. The power demand is in the order of microwatts per gate and this is a distinct advantage when designing large systems. CMOS chips are available in packages similar to those which house TTL. The available gate configurations are also similar. Since the input resistance seen at inputs to CMOS devices are very high, there is the possibility of a charge build-up at the inputs. Unless the inputs are terminated by a low resistance path, the electric fields associated with the accumulation of these charges can destroy the device. Due to this reason CMOS chips have to be handled with care and contained in the conductive foam packing until required for use.

Illustrative Examples

Example 9.1

A certain application requires using an op-amp in a non-inverting mode. A stage of 10 is necessary. If the feedback resistor has a value of 100 k, calculate the value of resistance that should be connected between the inverting input and the zero volt line to give the specified gain.

Solution

Using equation 9.2, the gain is given by

$$A = 1 + \frac{R_f}{R_1}$$

Rearranging this expression

$$R_1 = \frac{R_f}{A - 1}$$

Substituting given values

$$R_1 = \frac{100\ \text{k}}{10 - 1} = 11.1\ \text{k}\Omega$$

Example 9.2
The feedback network used in an op-amp circuit consists of a 0.1 μF capacitor and a 15 k resistor. The input is applied to the inverting input. Calculate the break frequency and comment on the shape of the response.

Solution
Using equation 9.3, the break frequency is given by

$$f = \frac{1}{2\pi\ c\ R_f} = \frac{1}{2\pi \times 0.02 \times 10^{-6} \times 15 \times 10^3}$$

$$= 530.5\ \text{Hz}$$

The response of this circuit will increase at 6 dB per octave below the break frequency. 6 dB per octave corresponds to 20 dB per decade.

Chapter 10

BASIC MEASUREMENTS

The operation of an electronic circuit can be verified by carrying out relevant measurements. These measurements always always concern some constant or variable parameter which characterizes the circuit. Using suitable measuring instruments we are able to quantify these parameters and thereby verify whether a circuit performs in accordance with a given design specification.

In any measurement situation it is desirable that we satisfy certain requirements associated with good measurement practice. First and foremost, we must ensure that the measuring instrument is suitable for the application. It should not unduly load the circuit or system under test. When a measuring instrument extracts too much power from the system being tested it is said to load the system. This loading effect is harmful in the way that it modifies the behaviour of the item under test, and whatever measurements we make then apply to the modified behaviour. If the loading effect is unavoidable it is important to know how and to what extent the system is being modified. We may then apply a correction factor to our results.

An aspect of measurement not often emphasized concerns the way in which the test item and the test instruments are arranged on the workbench. An approach based on good ergonomic practice can be of value in this context. It is desirable that such items as signal sources are placed to the left of the experimenter. Oscilloscopes and multimeters are best placed to the right of the experimenter. A convenient place for the power supplies would be immediately opposite.

The proper documentation of results is also very important when carrying out any measurement. The documentation should follow a logical pattern designed to present the information in a concise and clear manner.

Voltage Measurements

When carrying out voltage measurements we must make sure that the input impedance of the measuring instrument is high in comparison to the terminal impedance across which the measurement is made. It is good practice to ensure that the input impedance is greater at least by a factor of ten. The error due to the damping or loading effect may then be kept to within ten percent. When carrying out dc measurements, it is the resistive part of the impedance which is significant and needs to be considered. In ac voltage measurements, both the resistive and reactive parts of the impedance are significant. In general the reactive part will be due to the presence of capacitance. The voltage indicated by the instrument will be less than the actual value, if the reactive part of the input impedance compares with the resistive part. Since capacitive reactance is inversely proportional to frequency, it would be infinite at zero frequency. As the frequency is increased the reactance would drop and at some frequency it will be small enough to be comparable to the resistive part of the input impedance. This effect imposes a frequency limitation on the voltage being measured.

Current Measurements

Current measurements may be carried out either by connecting in series, a current measuring instrument and reading the current directly, or by taking a voltage measurement across a known resistance and calculating the current using Ohm's Law. If the first method is used we must make sure that the inclusion of the measuring instrument in series does not modify the current flowing in the circuit. Therefore ideally the instrument resistance must be zero. In practice it suffices if we ensure that the instrument resistance is small in comparison to the total circuit resistance. If the second method is used the voltage measuring instrument must have an input resistance which is high in comparison to the known resistance across which the measurement is made. If an additional resistance is included in the circuit for the purpose of obtaining the current, by measuring the voltage across it, we must make sure that the value chosen is small in comparison to the circuit resistance.

Power Measurements

The product of voltage and current gives the power in a dc circuit. The reader has been introduced to the relevant expressions in an earlier chapter. These expressions show that a knowledge of voltage or current is adequate for calculating power, if the resistance is known. This approach will be of most value to the hobbyist in the context of dc power measurements. AC power measurements are straightforward so long as the voltage and current are in phase. This implies that the impedance in question is purely resistive. The voltage across this impedance may be measured using an ac voltmeter or an oscilloscope. The rms power is obtained by dividing the square of the rms voltage by the resistance. The peak power is obtained by dividing the square of the peak voltage by the resistance. If the voltage and current are not in phase, the calculation of power requires a knowledge of the phase angle.

Two-port Network Measurements

A two-port network is one which has a pair of input terminals and pair of output terminals. Each pair of terminals is referred to as a port. There are several measurements we could carry out on this type of network. Gain, input resistance, output resistance and frequency response are a few quantities likely to be of interest to the hobbyist.

The gain or amplification of a two-port network is defined as the ratio of output voltage to input voltage. In this discussion we are concerned with signal voltages of sinusoidal waveshape. The diagram in Figure 10.1 shows the set-up for carrying out a gain measurement. The network is excited by a sinusoidal signal obtained from a suitable signal generator. A resistive pad is inserted between the generator and the network. This pad functions to terminate the input port by a low resistance, and attenuate the signal from the generator by a known amount. The signal voltage at the output port is measured using a suitable voltage measuring instrument. If the output voltage is greater than the input voltage the network behaves as an amplifier. If the output voltage is less than the input

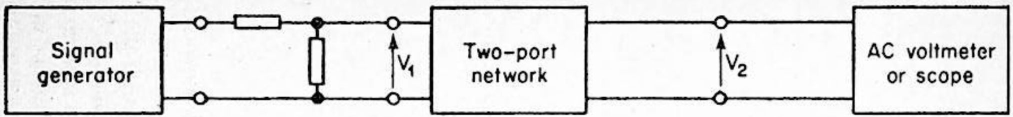


Fig. 10.1 Set-up for measurement of gain of two-port network.

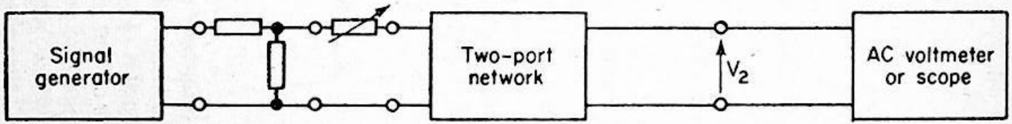


Fig. 10.2 Set-up for determining input resistance of two-port network.

Fig. 10.3 Set-up for measuring frequency response of two-port network.

voltage the network behaves as an attenuator. The gain is greater than unity for an amplifier, and less than unity for an attenuator. Gain is usually expressed in decibels. If the gain, expressed as a numeric is N, its value in decibels is given by $20 \log_{10} N$. Since the logarithm to base 10 of unity is zero, a gain of unity corresponds to 0 decibels.

The input resistance of a two-port network may be measured in the following manner. A signal is applied to the input-port and the output voltage noted down. A variable resistance is connected as shown in Figure 10.2 and its value adjusted until the output voltage is reduced by a factor of two. Assuming that the frequency of measurement is low enough to ignore reactive effects, the input resistance is given by the value of the variable resistance.

The output resistance may be found as follows. The input-port is excited by a signal adjusted to a suitable level. The output-port is terminated in turn by two known values of resistance. The voltage across these resistors are measured using a suitable voltage measuring instrument. The output resistance R_0 is then obtained employing the relationship given below.

$$R_0 = \frac{V_2 - V_1}{\dfrac{V_1}{R_1} - \dfrac{V_2}{R_2}} \qquad R_2 > R_1$$

The set-up for measuring the frequency response is shown in Figure 10.3. While maintaining the amplitude constant the frequency of the excitation signal is varied in steps over the specified frequency band. The output voltage is noted for each step. The results so obtained are presented on a graph on log-lin graph paper. The gain at one frequency is selected as the reference and the gain at all the other steps are related to this reference, and expressed in dBs. For an audio frequency amplifier the reference frequency is usually 1kHz.

Notes

BERNARDS No 228

Essential Theory for the Electronics Hobbyist

- In any hobby activity a background knowledge of the subject can considerably increase the enjoyment and satisfaction one derives from it. This view point applies, without any reservations whatsoever, to electronics.

- The object of this book is to supply the hobbyist with a background knowledge tailored to meet his or her specific requirements and the author has brought together the relevant material and presented it in a readable manner with minimum recourse to mathematics.

- Many formulae having a practical bearing are presented in this book and purpose designed examples are employed to illustrate their applications.

- An extremely useful addition to the library of any enthusiast.

BERNARDS (Publishers) LTD
The Grampians
Shepherds Bush Road
London W6 7NF
England

ISB

£1.25