UNCLASSIFIED

BASIC THEORY OF DIGITAL COMPUTERS

1 January 1957

VOLUME 2

IBM

UNCLASSIFIED

BASIC THEORY OF DIGITAL COMPUTERS

VOLUME 11

Contract No.

AF 30(635)-1404

1 January 1957

MILITARY PRODUCTS DIVISION
INTERNATIONAL BUSINESS MACHINES CORPORATION
KINGSTON, NEW YORK

BASIC THEORY OF DIGITAL COMPUTERS

DC1

PART 3

FUNDAMENTALS OF ELECTRICITY, MAGNETISM AND ELECTRONICS

Draft No. 2

INTERNATIONAL BUSINESS MACHINES CORPORATION
KINGSTON, NEW YORK

UNCLASSIFIED

PART 3

CHAPTER I

INTRODUCTION TO ELECTRICAL THEORY

1.1 MATTER AND ELECTRICITY

All matter is composed of mixtures and compounds of elements. For example, sea water is a mixture of water and various minerals, chief among which is salt. Salt is a chemical combination or compound of the two elements sodium and chlorine. Water,on the other hand, is a compound of the two elements hydrogen and oxygen. There are about one hundred known elements, each having unique properties.

The smallest possible sub-division of an element is the atom. Thus, an atom of one element is different in some way from the atoms of all other elements. For many years, it was thought that atoms were the basic building blocks of matter. However, it is now known that atoms are composed of various kinds of smaller particles. Moreover, it is known that atoms of different elements are composed of different numbers and arrangements of the same kinds of particles rather than of different kinds of particles.

From the point of view of electrical theory, the two important kinds of particles appearing in the structure of an atom are protons and electrons. All electrical phenomena depend upon the fact that protons attract electrons and vica versa,while protons repel other protons and electrons repel other electrons. There is an equality between protons and electrons; that is one proton and one electron satisfy each other. Thus, a grouping that includes the same number of protons as electrons attracts neither protons nor electrons and hence is said to be electrically neutral. By convention, a proton is considered to possess a positive charge while an elctron is considered to possess a negative charge. Thus, a grouping that includes more protons than electrons is said to be positively charged while a grouping that includes more electrons than protons is said to be negatively charged. A positive charge attracts protons and repels electrons.

The structure of an atom appears to be something like the structure of the solar system. Protons, together with other kinds of particles which need not be identified for the purposes of this discussion, form the nucleus of an atom. The electrons of an atom revolve about this nucleus in somewhat the same way that the earth and other planets revolve around the sun. Moreover, each of the electrons spins in the same way that the earth rotates about its own axis.

In general, atoms are electrically neutral by virtue of having an equal number of protons and electrons. However, it is possible for an atom to lose one or more of its electrons, in which case it possesses a positive charge and is said to be a positive ion. It might be expected that an atom could become negatively charged by losing one or more protons. However, this is not the case, since the position of the protons in the nucleus of the atom makes their escape impossible. On the other hand, an atom can become negatively

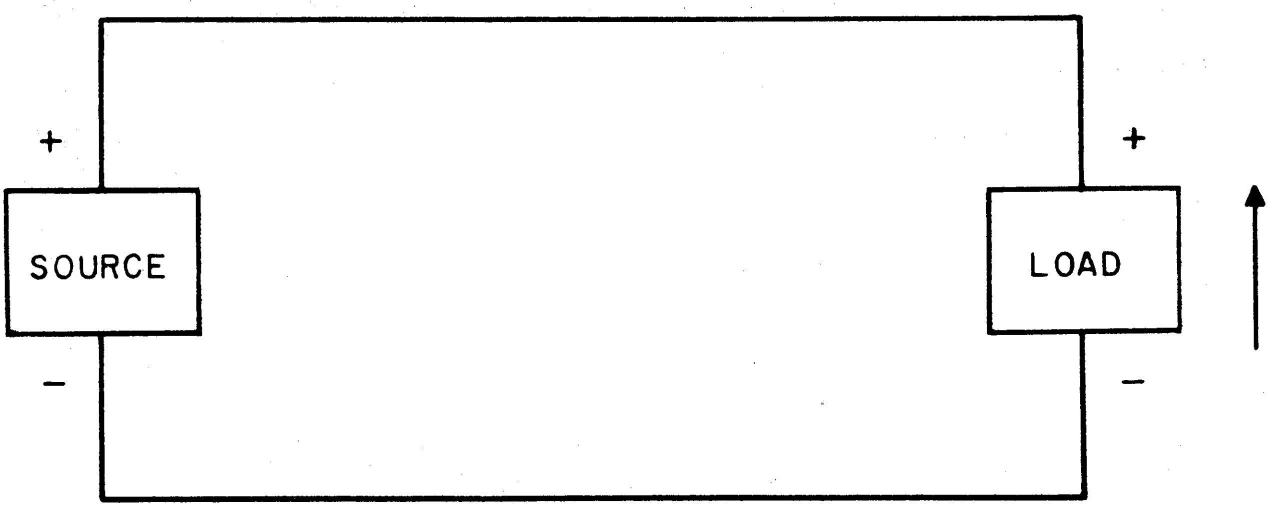


Figure 3-1

charged by gaining excess electrons, in which case it is said to be a negative ion.

1.2 ELECTRICAL CIRCUITS

Since electrons belonging to a particular atom are bound to that atom by mutual attraction between themselves and corresponding protons in the nucleus, work must be done in order to dislodge them from the atom. Thus some kind of force must be applied.

An electric circuit generally comprises a source which supplies the force to move electrons and a load through which electrons are moved. The source and the load are connected so as to form a closed path as illustrated in Figure 3-1. The source functions to create a deficiency of electrons at one of its terminals and an excess of electrons at the other. Thus, electrons from the load are attracted to the terminal of the source which is deficient in electrons, while electrons in the load are repelled from the terminal of the source which has an excess of electrons. The result is a motion of electrons around the circuit in the direction indicated by the arrow in Figure 3-1. This motion is called electrical current. The force supplied by the source is called an emf (or electro-motive force).

As has already been stated, it is not possible for protons to escape from their respective atoms. Thus, in general, electrical current is characterized by the motion of electrons. However, under certain circumstances, ions (i.e. atoms having either an excess or a deficiency of electrons) move through a circuit. This type of motion of charged particles is called electrolytic current to distinguish it from current characterized by the motion of electrons.

1.3 CONDUCTORS AND INSULATORS

Before electrons can be moved around a circuit, they must first be detached from their respective atoms. In some elements; one, two or three of the outer electrons of each atom are rather loosely bound to that atom and thus can be detached from it by the application of a relatively small force. Such elements are said to be good conductors because electric current can be caused to flow through them relatively easily. In other elements, all the electrons of each atom are bound tightly to that atom so that a relatively large force is required to detach an electron from an atom. Such elements are said to be good insulators because a relatively large force is required to cause current to flow through them. There is no such thing as a perfect insulator because there is no element whose atoms will not give up electrons if a sufficiently large force is applied. On the other hand, there is no such thing as a perfect conductor because there is no element whose atoms will give up electrons without the application of any force. However, metals such as copper, silver and gold are very good conductors while substances such as glass and rubber are very good insulators.

1.4 BATTERIES

Sone source of emf is a cell which converts chemical energy into electrical energy. The action of such a cell depends upon the fact that certain compounds tend to separate into more simple compounds or elements when dissolved in water. In the course of this separation positive and negative ions of the more simple compounds or elements are produced. A solution containing these ions is called an electrolyte. Two rods or plates called electrodes are suspended in the electrolyte. One of the electrodes is composed of some substance which tends to give up electrons to the positive ions while the other electrode is composed of some substance which tends to take electrons from the negative ions. Thus, one electrode becomes deficient in electrons while the other gains an excess of electrons so that a potential difference is said to exist between the two.

A cell such as just described can be made using dilute sulphuric acid (H_2SO_4) and zinc and copper plates as electrodes. The sulphuric acid separates into positive hydrogen ions and negative SO_4 ions. The negative ions combine with the zinc to form zinc sulphate, excess electrons being left on the zinc electrode as a by-product of the reaction. At the same time, the copper electrode gives up electrons to the positive hydrogen ions. When a certain potential difference exists between the two electrodes, the chemical action stops; however if the cell is connected as the source in a closed circuit such as that of Figure 3-1, then current is drawn through the load and the chemical action is continuous.

It should be understood that chemical energy is expended in driving current through a circuit so that the process

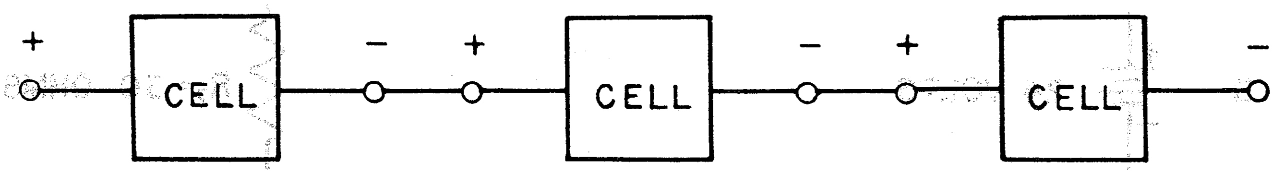


Figure 3-2

cannot continue indefinitely. What happens is that the electrolyte or one of the electrodes is used up. In some cells, the chemical process which generates the potential difference is reversable, so that the cell can be recharged by driving current through it in a direction opposite to that occurring in normal operation. Such cells are called secondary cells. In other cells, such as the one just described, the process is not reversable so that when the electrolyte or electrode has been used up, the cell must be discarded. Such cells are called primary cells.

As already stated, a cell reaches chemical equilibrium and the chemical action stops when a certain potential difference exists between the electrodes. The magnitude of this potential difference depends upon the particular electrolyte and electrodes used. On the other hand, the amount of current that can be drawn from the cell and the length of time that it will last depend upon the dimensions of the electrodes and of the cell.

When potential differences higher than those which can be obtained from a single cell are desired, several cells can be connected in series as shown in Figure 3-2. Such a combination of cells is called a battery.

1.5 OHM'S LAW

As has already been stated, current is the motion of electrons through a closed circuit which occurs in response to the application of an electromotive force (emf) to that circuit. The ratio of the emf applied to a particular circuit to the resultant current that flows through that

TABLE 3-1

UNITS AND SYMBOLS FOR EXPRESSING VALUES OF EMF, CURRENT AND RESISTANCE

Current = I

amperes = A

milliamperes = amperes X 10^{-3} = MA

Microamperes = amperes X 10^{-6} = μA

Electromotive force = E

volts = V

millivolts = volts X 10^{-3} = MV

Resistance = R

ohms = Ω

kilohms = X 10^{3} = K

megohms = X 10^{6} = M

circuit is called the resistance of the circuit. This ratio, known as Ohm's Law, is written:

$$\frac{E}{I} = R \qquad (1)$$

where:

E=applied emf in volts

I=current in amperes

R=resistance in ohms

An ampere of current is said to be flowing in a circuit when the motion of electrons past some reference point in the circuit is at the rate of 1 coulomb per second, where a coulomb is 6.3×10^{18} electrons.

The unit of emf, that is the volt, is based upon the potential difference (or voltage) generated by a standard cell. The unit of resistance, that is the ohm, is defined in terms of volts and amperes as follows:

If an emf of 1 volt drives a current of 1 ampere through a circuit, then the resistance of that circuit is 1 ohm.

For convenience in expressing values of emf, current or resistance which are either very large or very small when stated in the fundamental units, other units are used. Each of these units is equivalent to the fundamental unit times some power of ten. Units in common use are listed in Table 3-1.

1.6 DISTRIBUTED AND LUMPED PARAMETERS

The term circuit is extremely generalized. It can refer to a current path or portion of a current path or it can refer to a set of current paths belonging to some single device such as a radio. For example, it is usual to speak of all those wall outlets and lamps which receive current through the same fuse in a household electrical system as belonging to the same circuit. On the other hand, the complete path between power plant and home electrical system is also a circuit.

The purpose of a portion of any circuit is simply the transmission of electrical power from the source of emf to the physical location where it is to be used to produce some such end product as light, heat, sound or motion.

The materials used to form the transmission path portion of any circuit should offer a minimum of opposition to the flow of current. Thus materials such as copper which are good conductors are used. However, as has already been noted, there is no such thing as a perfect conductor. For example, copper has a resistivity of 10.4 ohm-cm per foot, which means that a speciman of copper 1 foot long having a cross-sectional area of 1 circular mil offers a resistance of 10.4 ohms to the flow of current.

The resistance of a transmission path consisting of wire is distributed evenly along the wire. Thus, in the case of wire, resistance is said to be a distributed parameter.

There are many instances in which the purposes of a circuit are served by concentrating resistances at particular points. For example, by driving a current of I amperes through a resistance or R ohms, a signal of V volts can be caused to appear across that resistance (i.e. a difference of potential of V volts can be created between one end of the resistance and the other. Elements called resistors, which concentrate various quantities of resistance in small volumes, are one of the most common of circuit elements. The resistance of a resistor is said to be a lumped parameter, since it appears at one place in the circuit rather than being distributed evenly over some significant physical portion of the circuit.

1.7 SERIES AND PARALLEL CIRCUITS

A series circuit is one which offers but one path for current flow while a parallel circuit is one which offers more than one path. A circuit which offers a single path over portions of its length and multiple paths over other portions of its length is a series-parallel circuit.

1.7.1. Total Resistance of Series Circuit

The total resistance of a series circuit is equal to the sum of the individual resistances. Thus for a circuit comprising N resistances in series, the total resistance R_t is given by the equation:

$$R_t = R_1 + R_2 \ldots\ldots + R_n \qquad (2)$$

In the circuit of Figure 3-3, for example, battery B provides an emf of 24 volts which drives a current I through series-connected resistors R_1 and R_2. The magnitude of current I is, by Ohm's Law:

$$I = \frac{E}{R_t}$$

where

$R_t =$ total resistance $= R_1 + R_2 = 10$ ohms $+ 30$ ohms $= 40$ ohms

thus:

$$I = \frac{24}{40} = 0.6 \text{ amperes}$$

Now that the current is known, the voltage drop through each of the resistors can be found by a further application of Ohm's Law as follows:

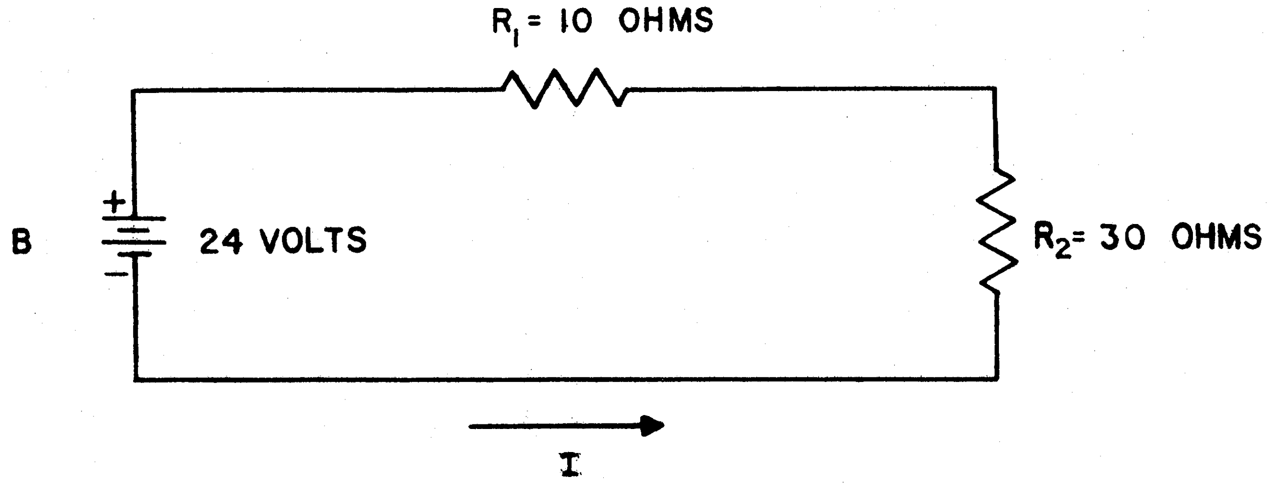


Figure 3-3

$V_1 = IR_1 = (0.6)\ (10) = 6$ volts

where:

$V_1 =$ voltage drop through R_1

Moreover:

$V_2 = IR_2\ (0.6)\ (30) = 18$ volts

where:

$V_2 =$ voltage drop through R_2

Notice that the sum of the voltage drops through the load is equal to the emf supplied by the source, that is:

$E = V_1 + V_2$

24 volts $=$ 6 volts $+$ 18 volts

In this case the internal resistance of the source is considered to be negligible, that is the source is assumed to provide a virtually constant emf regardless of the amount of current that is drawn from it. If E is the potential difference between the terminals of a source when no current is being drawn from that source and if V is the potential difference when a current I is being drawn, then the internal resistance, R_i, of the source, is

$$R_i = \frac{E-V}{I}$$

1.7.2 Voltage Dividers

The total voltage drop through a series load is divided among the individual resistors comprising that load in proportion to the magnitudes of their individual resistances. Thus, for the circuit of Figure 3-3:

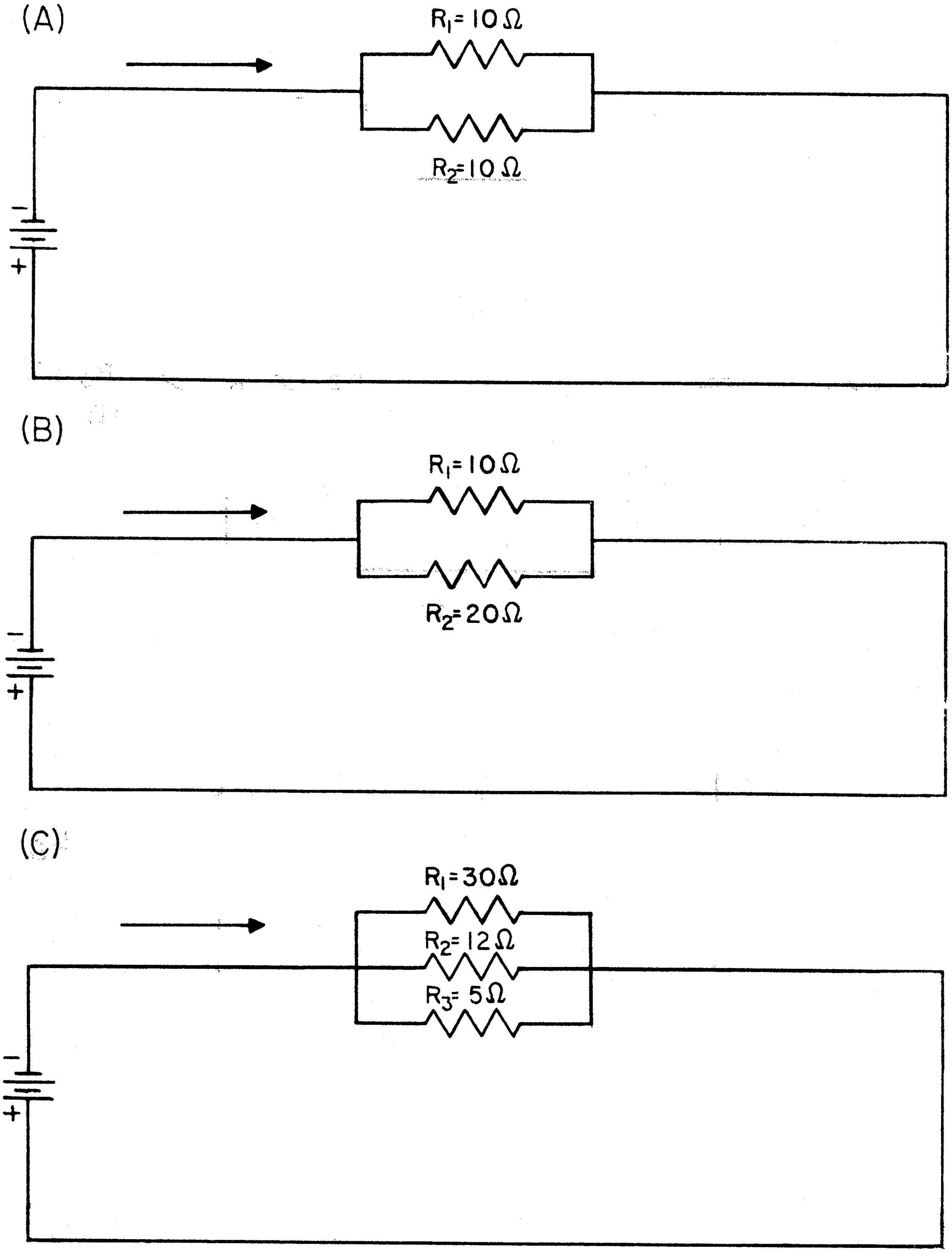


Figure 3-4

$$V_1 = V_t \left(\frac{R_1}{R_1 + R_2} \right) = 24 \left(\frac{10}{40} \right) = 6 \text{ volts}$$

$$V_2 = V_t \left(\frac{R_2}{R_1 + R_2} \right) = 24 \left(\frac{30}{40} \right) = 18 \text{ volts}$$

where V_t, the total voltage drop through the load, equals E, the applied emf, by virtue of the assumption of negligible source resistance.

The general equation for the voltage drop, V_1 through any single resistor, R, connected in a series circuit of total resistance R_t, where V_t is the total drop across R_t is:

$$V_1 = V_t \left(\frac{R}{R_t} \right) \qquad (3)$$

The arrow in Figure 3-3 shows the direction of current flow which is assumed throughout this book. This direction corresponds to the motion of electrons through the load from the negative terminal of the battery to the positive terminal of the battery. Before the phenomenon of current was understood in terms of electron flow; the opposite convention was adopted; that is current was assumed to flow out of the positive terminal of a battery and into the negative terminal. Since both conventions are in use today, it is important, when approaching a theoretical discussion in which the direction of the current is significant, to establish which current convention is being used.

1.7.3 Total Resistance of Parallel Circuit

Three parallelcircuits are shown in Figure 3-4. Notice that the full emf of the source is applied across each of the resistors of a parallel circuit. Thus, by Ohm's Law, the currents through the various resistors of a parallel circuit are given by equations of the form:

$$I_1 = \frac{E}{R_1},\quad I_2 = \frac{E}{R_2}, \ldots, I_n = \frac{E}{R_n} \qquad (4)$$

The total current drawn from the source, I_t, is simply the sum of the currents drawn through the individual resistors. Thus:

$$I_t = I_1 + I_2 + \ldots\ldots + I_r \qquad (5)$$

Moreover, the total resistance of the circuit, R_t, is, by Ohm's Law:

$$R_t = E/I_t \qquad (6)$$

Substituting equations (4) into equation (5) yields:

$$I_t = \frac{E}{R_1} + \frac{E}{R_2} + \ldots\ldots + \frac{E}{R_n} \qquad (7)$$

Substituting equation (6) into equation (7) yields:

$$R_t = \frac{E}{\frac{E}{R_1} + \frac{E}{R_2} + \ldots\ldots + \frac{E}{R_n}} \qquad (8)$$

Dividing the numerator and denominator by E in equation (8) yields:

$$R_t = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \ldots\ldots \frac{1}{R_n}} \qquad (9)$$

which is the general equation for the total resistance in a parallel circuit.

As an example, the total resistance of the circuit of Figure 3-4 (a) is found as follows:

By equation (9)

$$R_t = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$$

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

$$= \frac{R_1 R_2}{R_2 + R_1} \qquad (10)$$

Since, in the example,

$$R_1 = R_2 = R = 10 \text{ ohms}$$

equation (10) reduces to:

$$R_t = \frac{R^2}{2R} = \frac{R}{2} = \frac{100}{20} = 5 \text{ ohms}$$

To generalize this result, the total resistance, R_t of any parallel circuit formed of n resistors of resistance R is given by the equation:

$$R_t = \frac{R}{n} \qquad (11)$$

Thus, for, example, the total resistance of 1000-ohm resistors connected in parallel is:

$$R_t = \frac{1000}{10} = 100 \text{ ohms}$$

Consider now the example of Figure 3-4-(b). Here, the parallel circuit comprises two resistors of different values so that equation (11) cannot be applied. However, for the special case of two resistors, equation (9) is written as follows:

$$R_t = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$$

$$= \frac{1}{\frac{R_2 + R_1}{R_1 \quad R_2}}$$

$$= \frac{R_1 \quad R_2}{R_1 + R_2} \qquad (12)$$

Substituting the values of R_1 and R_2 from the example of Figure 3-4 (b) into equation (12):

$$R_t = \frac{(10)\ (20)}{10 + 20} = \frac{200}{30} = \frac{20}{3} \text{ Ohms}$$

As an example of the direct substitution of numerical values into general equation (9), consider Figure 3-4 (c). Here:

$R_1 = 30$ ohms $R_2 = 12$ ohms $R_3 = 5$ ohms

Thus:

$$R_t = \frac{1}{\frac{1}{30} + \frac{1}{12} + \frac{1}{5}}$$

$$= \frac{1}{\frac{2 + 5 + 12}{60}}$$

$$= \frac{1}{\frac{19}{60}}$$

$$= \frac{60}{19} \text{ ohms}$$

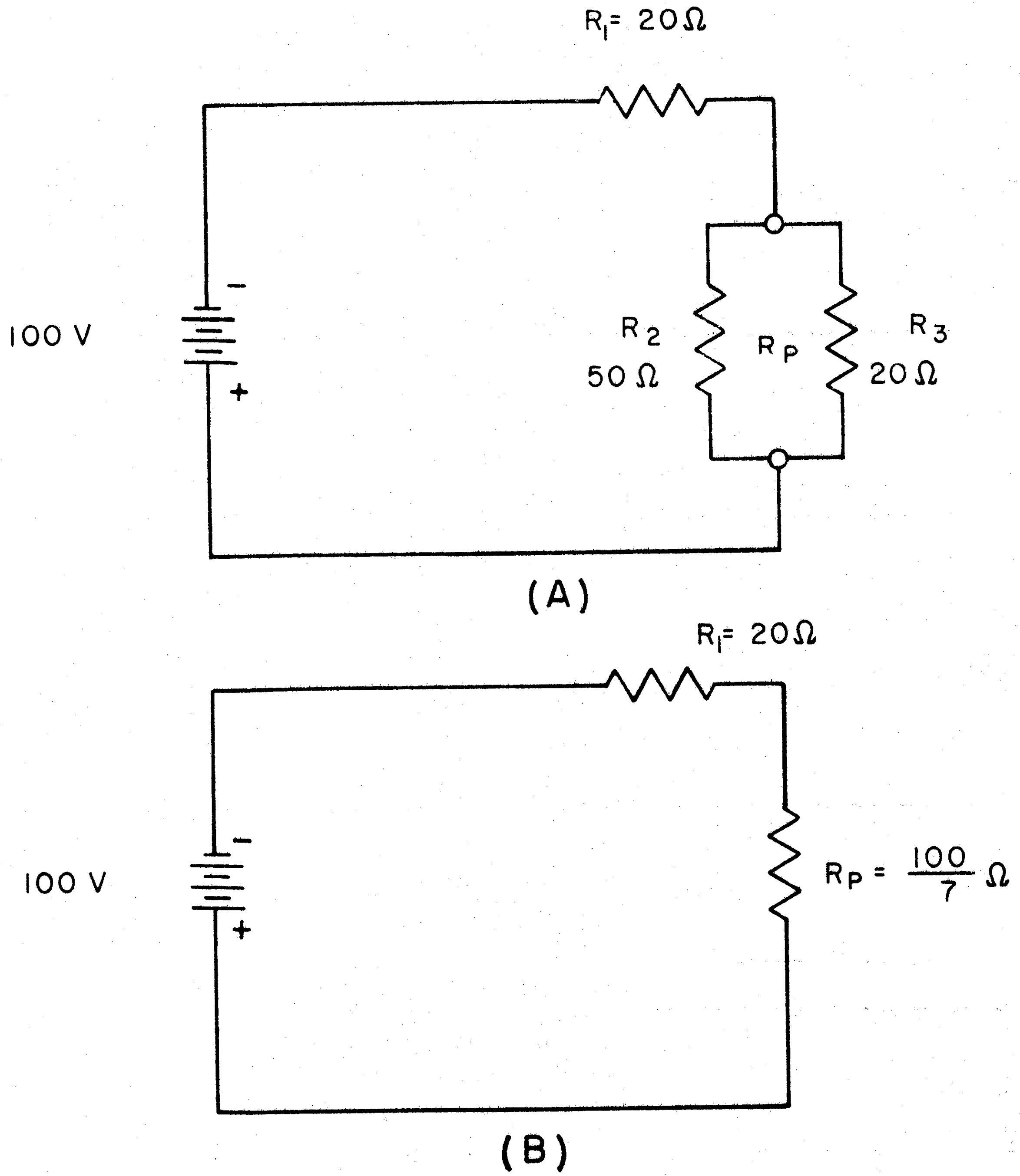


Figure 3-5

To summarize the findings of this section:

The total resistance, R_t, of n resistors connected in parallel is given by the general equation:

$$R_t = \frac{1}{1/R_1 + 1/R_2 + \ldots . 1/R_n} \tag{9}$$

If all n resistors are of the same value, this equation reduces to:

$$R_t = \frac{R}{n} \tag{11}$$

If only two resistors are involved but these are of different values, the general equation can immediately be written:

$$R_t = \frac{R_1 R_2}{R_1 + R_2} \tag{12}$$

1.7.4 Series-Parallel Circuits

In order to determine the total resistance of a series-parallel circuit:

a. Solve for the total resistance of each parallel portion of the circuit.

b. Solve for the resistance of the resulting equivalent series circuit.

As an example, consider the series-parallel circuit of Figure 3-5-(a).

The total resistance, R_p, of the parallel combination of resistors R_2 and R_3 is found by equation (11) as follows:

$$R_p = \frac{(50)\ (20)}{50 + 20}$$

$$= \frac{1000}{70} = \frac{100}{7} \text{ ohms}$$

Since R_p is in series with R_1, as shown in Figure 3-5 (b), the total resistance of the circuit, R_t, is: by equation (2)

$$R_t = R_1 + R_2$$

$$= 20 + \frac{100}{7}$$

$$= \frac{240}{7} \text{ ohms}$$

The total current through the circuit is, by Ohm's Law:

$$I_t = \frac{E}{R_t}$$

$$= \frac{100}{240/7}$$

$$= \frac{70}{24} = \frac{35}{12} \text{ ampere}$$

The voltage drop across R_p as given by equation (3) is:

$$V_p = V_t \left(\frac{R_p}{R_1 + R_p} \right)$$

$$= 100 \left(\frac{100/7}{240/7} \right)$$

$$= \frac{125}{3} \text{ volts}$$

Here, the assumption is made that the internal resistance of the source is negligible so that $V_t = E = 100$ V.

The voltage drop across R_1 is found similarly:

$$V_1 = V_t \left(\frac{R_1}{R_1 + R_p} \right)$$

$$= \left(\frac{20}{240/7} \right)$$

$$= \frac{775}{3} \text{ volts}$$

The total voltage drop across $R_1 + R_p$ must equal the sum of the individual voltage drops. Thus:

$$V_t = V_1 + V_p$$

$$= \frac{175}{3} + \frac{125}{3}$$

$$= \frac{300}{3}$$

$$= 100 \text{ volts}$$

which checks.

The current through R_1 is simply the total current I_t. The currents through R_2 and R_3 can be found by substituting the values of these resistors and the voltage drop across their equivalent resistance, R_p (which is R_p), into Ohm's Law as follows:

$$I_2 = \frac{V_p}{R_2}$$

$$= \frac{\frac{125}{3}}{50}$$

$$= \frac{5}{6} \text{ ampere}$$

$$I_3 = \frac{V_p}{R_3}$$

$$= \frac{\frac{125}{3}}{20} \qquad = \frac{25}{12} \text{ amperes}$$

The total current in the circuit, I_t, must equal the sum of the currents through resistors R_2 and R_3. Thus:

$$I_t = I_2 + I_3$$

$$= \frac{5}{6} + \frac{25}{12}$$

$$= \frac{35}{12} \text{ amperes}$$

which checks.

1.8 POWER

When mass is raised against the force of gravity, work is done on that mass. For example, work must be done in pumping water out of a well. On the other hand, when mass is allowed to fall with the force of gravity, that mass does work. Thus water running down a millrace turns a water wheel.

The electrical analogy to the motion of mass against or with the force of gravity involves the motion of electrons away from or toward a positive charge. Thus an electrical source does work in moving electrons from a positively charged terminal to a negatively charged terminal. On the other hand, electrons moving through an electrical load (from the negative terminal of the source toward the positive terminal of the source) perform work on that electrical load.

Power is defined as rate of doing work. Thus, electrical power is the product of the potential difference through which electrons are moved and the rate at which they are moved through that potential difference. Moreover, the potential difference between the terminals of a source is by definition the emf, E, of that source while the rate of motion of electrons is the electrical current, I. Thus the power supplied by a source of emf is given by the equation:

$$P = EI \qquad (13)$$

The unit of electrical power is the watt. When a current of 1 ampere is driven through a potential difference of 1 volt, then 1 watt of power is developed.

A second expression for power is:

$$P = I^2R \qquad (14)$$

Equation (14) is derived from equation (13) as follows:

Ohm's Law is written as an explicit expression of emf:

$$E = IR$$

Equation (15) is substituted in equation (13), yielding equation (14), that is

$$P = (IR) = I^2R$$

As an example of the power characteristics of a circuit, consider the series-parallel circuit of Figure 3-5. The power developed in this circuit is:

$$P = EI_t$$

where E 100 volts as noted in the figure

and I_t 35/12 ampere as found in Section 1.7.4 above.

Thus:

$$P = (100)\,(35/12)$$

$$= \frac{875}{3} \text{ watts}$$

Since the voltage drop, V_1, through R_1 is known, the power dissipated through R_1 can also be found by means of equation (13) as follows:

$$
\begin{aligned}
P_1 &= V_1 I_t \\
&= \left(\frac{175}{3}\right)\left(\frac{35}{12}\right) \\
&= \frac{6125}{36} \text{ watts}
\end{aligned}
$$

which checks with the result obtained by application of equation (14).

Since both the voltage drop through R_2 and R_3 and the currents through each of these resistors have already been found, either equation (13) or equation (14) can be used to find the power dissipated through each of these resistors. Applying equation (14):

$$
\begin{aligned}
P_2 &= \left(\frac{5}{6}\right)^2 (50) \\
&= \frac{625}{18} \text{ watts}
\end{aligned}
$$

and:

$$
\begin{aligned}
P_3 &= \left(\frac{25}{12}\right)^2 (20) \\
&= \frac{3125}{36} \text{ watts}
\end{aligned}
$$

The total power dissipated through all three resistors is:

$$
\begin{aligned}
P_t &= P_1 + P_2 + P_3 \\
&= \frac{6125}{36} + \frac{625}{18} + \frac{3125}{36} \\
&= \frac{10500}{36} \\
&= \frac{875}{3} \text{ watts}
\end{aligned}
$$

which is equal to P, the power developed by the source. This corresponds to the fact that the internal resistance of the source is assumed to be negligible. In an actual source, some of the power developed is dissipated through the internal source resistance.

The power dissipated through the resistance appears in the form of heat.

PART 3

CHAPTER 2

MAGNETISM

2.1 INTRODUCTION

Although magnetism is a property which has been known for many centuries, it is, relatively speaking, only recently that its relationship to electricity has been noted. Since magnetism is involved in the operation of practically all electrical apparatus a thorough understanding of its principles is essential. Materials that possess this property of magnetism are known as magnets and are further subdivided into natural magnets, permanent magnets, and electromagnets.

2.2 NATURAL MAGNETS

Natural magnets are composed of iron ore, known as magnetite, which naturally possesses the property of magnetism. For many centuries these natural magnets were used for navigation by the ancients, because of their property of pointing north and south when suspended freely. It was through this particular use that the natural magnet acquired its name of lodestone or leading stone. These natural magnets also have the ability of imparting magnetism to a piece of iron or steel solely by rubbing the materials together.

2.3 PERMANENT MAGNETS

A permanent magnet is a piece of hard steel, steel alloy, or iron, which has been magnetized either by the

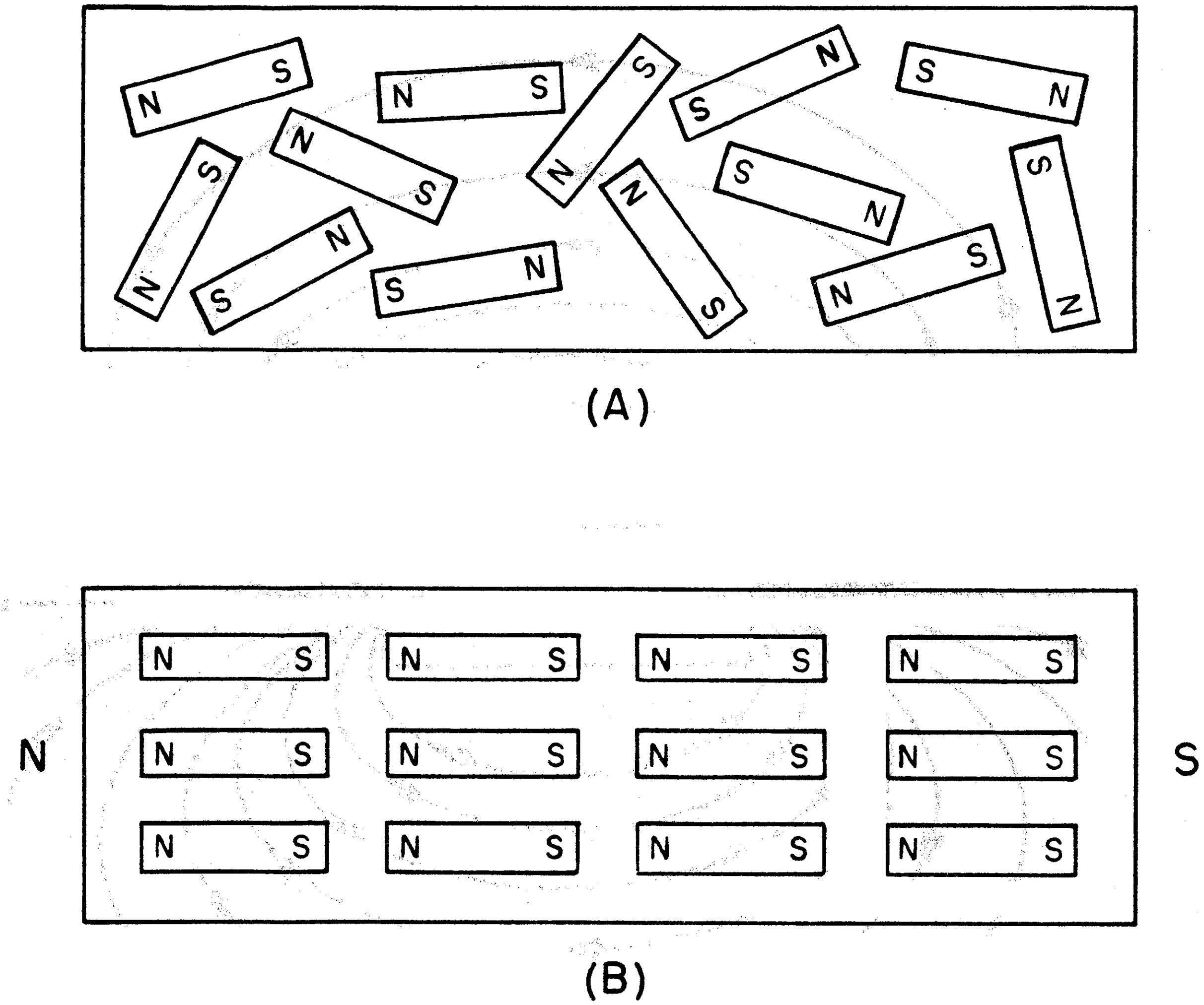


Figure 3-6

influence of another magnet (as mentioned above) or by means of an electrical current.

Under normal conditions a hard steel bar when magnetized retains its magnetism for a long period of time. However, if it is subjected to extreme heat or jarring its magnetism may be destroyed. Soft iron, on the other hand, is much more readily magnetized than hard steel, but tends to lose its magnetism within a relatively short time. Therefore, hard steel or its alloys are usually used as permanent magnets whereas iron or its alloys are used as cores in electromagnets or in situations where this ability to readily lose and regain magnetism is advantageous (such as in computer storage systems).

What condition exists in a piece of steel bar after it is magnetized that is not present in its natural state? This question has been answered by many theories. One of the most useful theories that has been advanced is the Weber and Ewing Theory. This theory states that the molecules which make up the magnetic material are themselves small magnets each containing a north and south pole. Before the magnetizing force is applied (unmagnetized state) these molecules are arranged in a haphazard manner, as shown in Figure 3-6(A). In this state, since the various molecular north and south poles neutralize each other no external magnetic effect is produced. However, when a magnetizing force is applied, each of the small magnets tend to arrange themselves parallel to each other and their north poles all point in the same general direction as

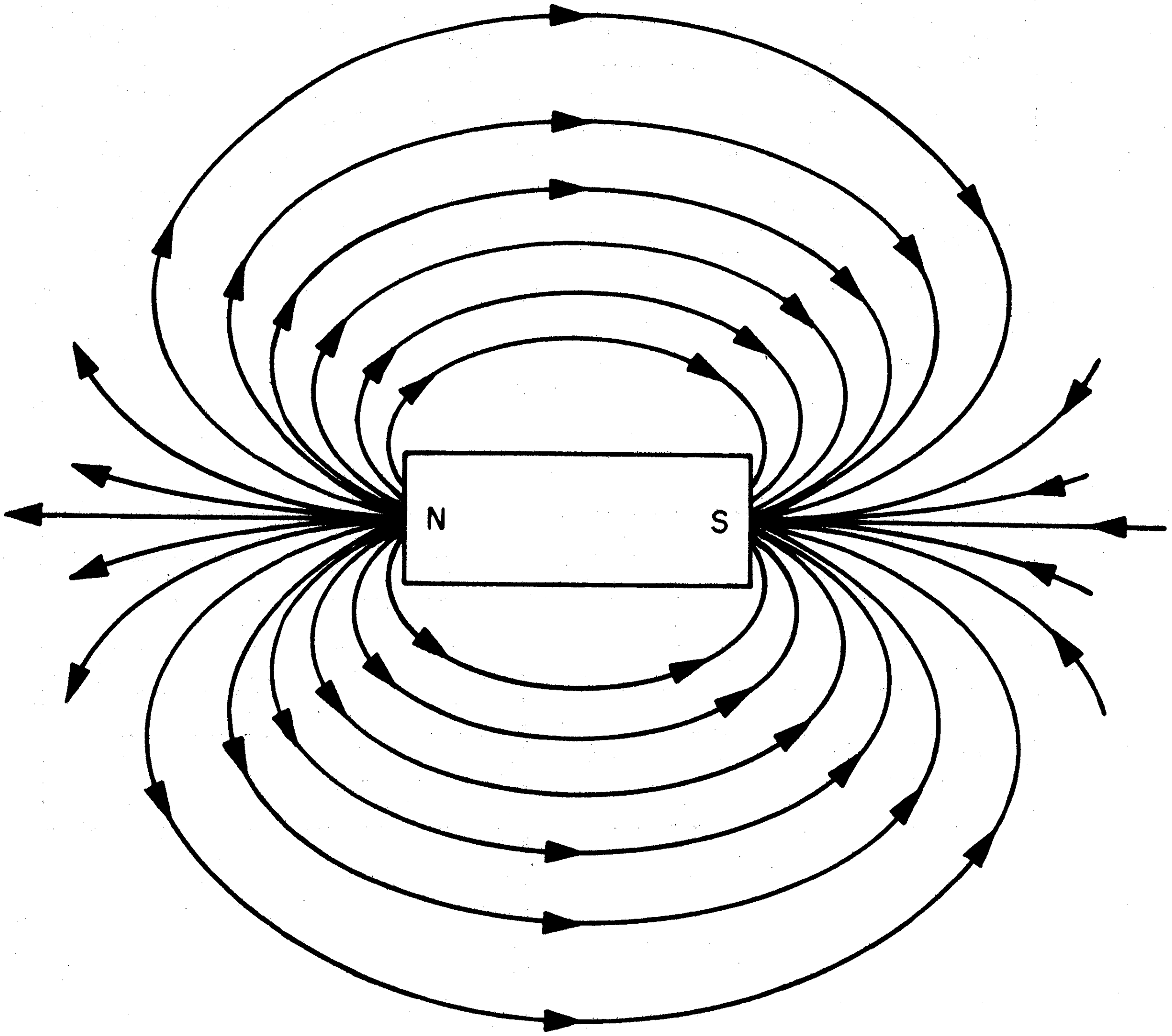


Figure 3-7

the magnetizing force, as shown in Figure 3-6 (B). This theory further goes to explain that if the permanent magnet is ground up into small particles, each of the particles possesses the properties of a bar magnet, each particle maintaining its own north and south pole. The amount of alignment of these molecules naturally depends on such factors as the magnitude and direction of the magnetizing force and the composition of the material to be magnetized. It is with the use of this theory that phenomena such as the saturation curve (B-H) and hysteresis can be explained. These phenomena are discussed in section 2.9 and section 2.10, respectively.

Magnetism itself is further defined as the force which surrounds any region that is magnetic in nature. It is represented by lines that emerge from one pole (north) of a magnet and enter into the other pole (south), and are called the magnetic lines of force or lines of induction as shown in Figure 3-7. These magnetic lines of force are also assumed to pass through the magnet from the south pole to the north pole. The path through which these magnetic lines of force pass is known as the magnetic circuit. These magnetic lines of force are also able to exert influence on any piece of iron or electric currents that are located in this region. This region is known as the magnetic field and its strength is represented by the density of these lines at any particular point.

2.4 ELECTROMAGNETISM

The relationship between magnetism and electricity

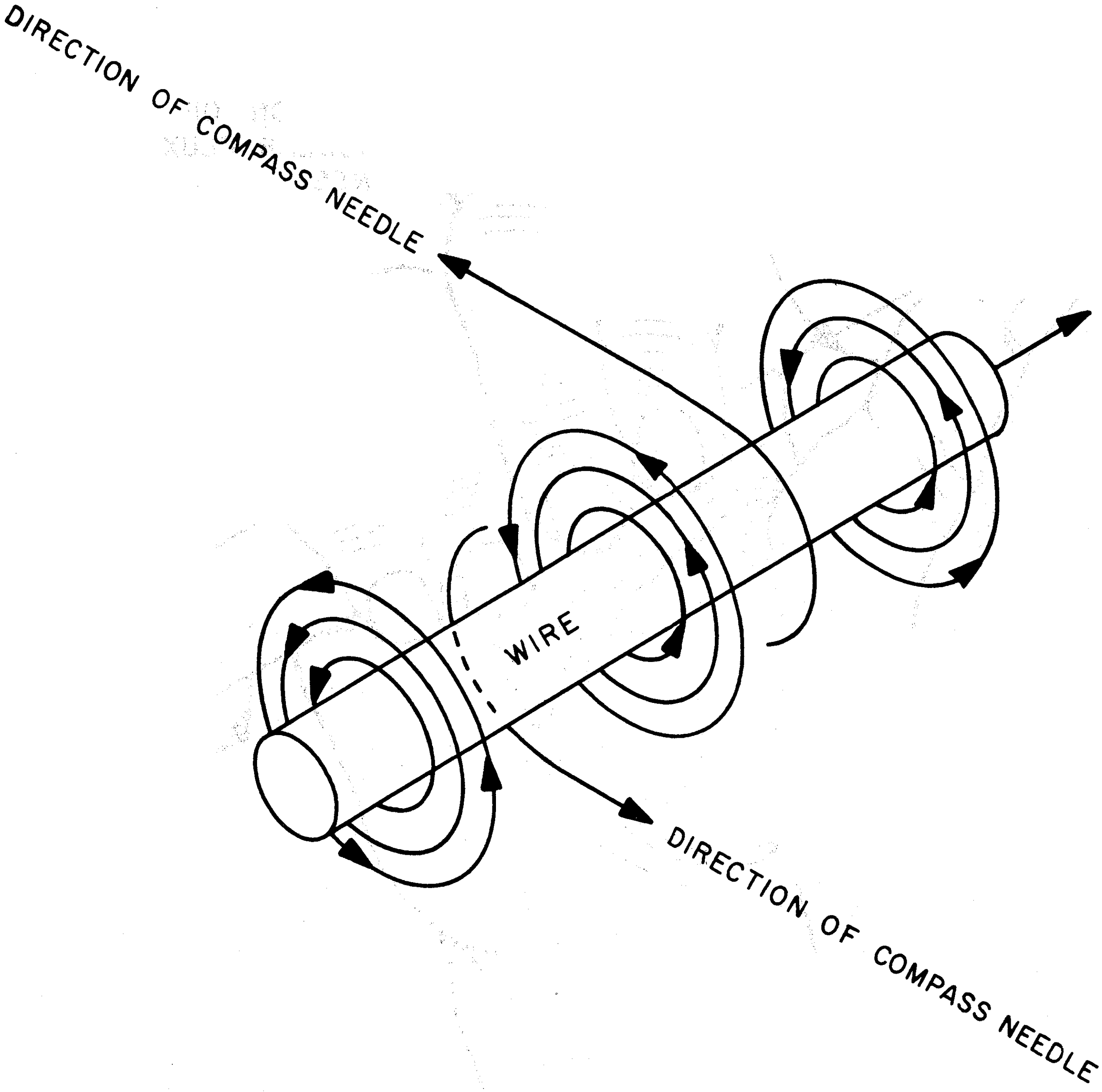

DIRECTION OF COMPASS NEEDLE
WIRE
DIRECTION OF COMPASS NEEDLE

Figure 3-8

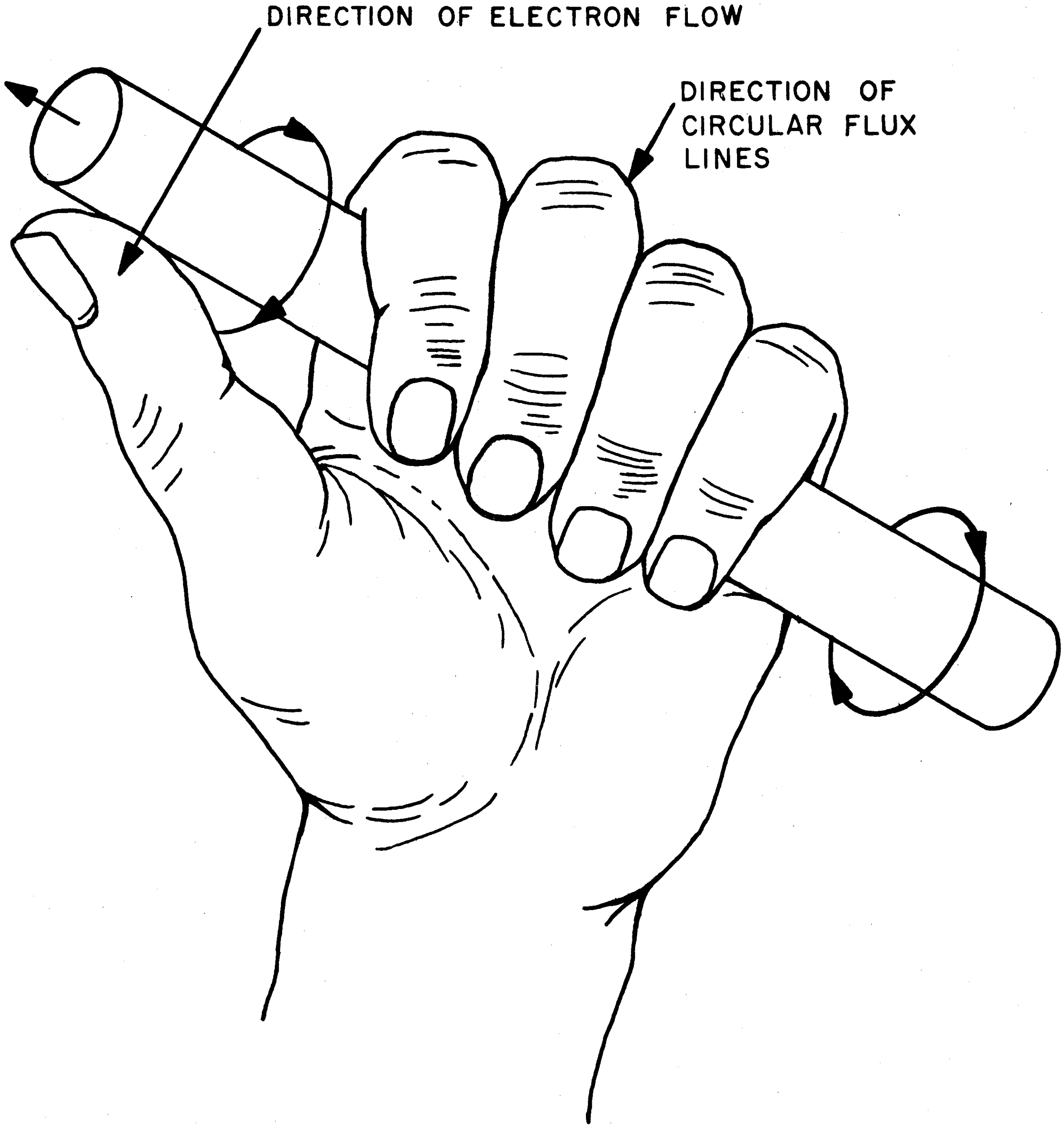

DIRECTION OF ELECTRON FLOW
DIRECTION OF CIRCULAR FLUX LINES

Figure 3-9

is demonstrated by passing a current through a conductor to show that a magnetic field is produced. A simple experiment demonstrates this fact. If a current is passed through a wire as shown in Figure 3-8 certain effects are noted in the surrounding space. When a compass needle is placed over the wire it aligns itself in a position perpendicular to the wire in the direction shown, whereas if it is placed below the wire, it arranges itself perpendicular to the wire facing in the opposite direction. In all cases it aligns itself perpendicular to the wire. Therefore, it is evident that the current produces some sort of a force in the region that acts in a circular direction surrounding it. That force is magnetism, which contains the same characteristics as the magnetism associated with permanent magnets. If the left hand is held as shown in Figure 3-9 so that the thumb points in the direction of the electron flow, then the remaining fingers indicate the direction of the circular whirls of magnetic flux around the wire.

2.5 PROPERTIES OF MAGNETIC LINES OF FORCE

Magnets, whether of the natural, permanent or electromagnetic type have certain properties which are attributed to the magnetic lines of force. These properties are as follows:

a. Relation to current flow: The magnetic lines of force,also known as magnetic flux,are set up by moving electric fields, and always surround these fields in the form of closed loops. The lines that normally emanate from or terminate on an electrically charged particle are

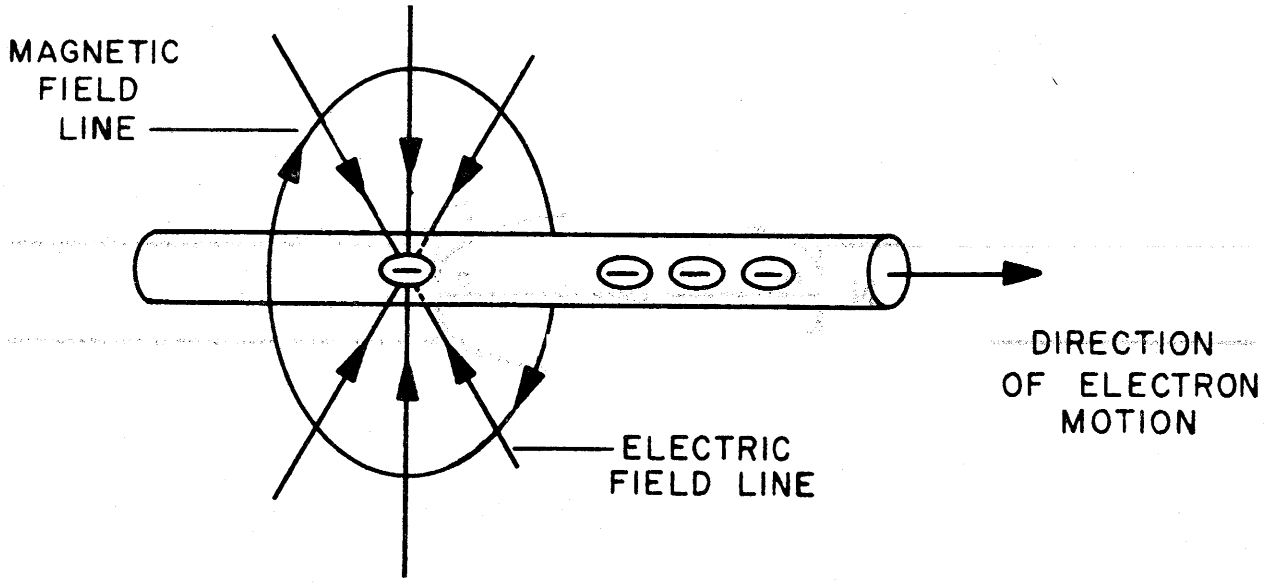

MAGNETIC FIELD LINE
ELECTRIC FIELD LINE
DIRECTION OF ELECTRON MOTION

Figure 3-10

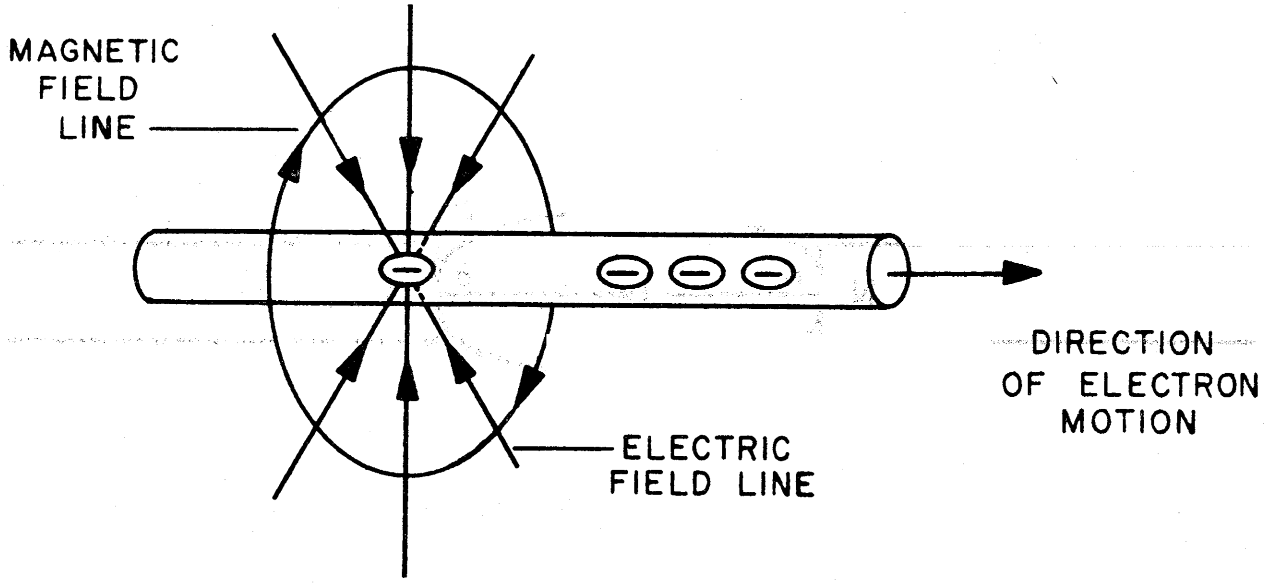

MAGNETIC FIELD LINE
DIRECTION OF ELECTRON MOTION
ELECTRIC FIELD LINE

Figure 3-10

the electric field lines. By convention, a positive charge is one from which the field lines emanate and a negative charge is one in which the field lines from some other positive charge terminate.

The associated field lines move when the charged body moves, which constitues a current flow. This current flow immediately produces loops of magnetic flux that enclose the moving charges. This is illustrated in Figure 3-10. Some of the free electrons of a conductor are shown. The radial spokes are formed by the electric field lines which are directed toward the electrons. Only a few of the radial field lines for one electron are shown for the sake of clarity. Since a charge and its associated field move as a unit either one can be considered as producing the magnetic field by virtue of its motion.

b. Tension in lines: The second property exhibited by magnetic lines of force is that of tension. The lines behave as though they are stretched rubber bands. If the magnetic lines of force extend from the face of one piece of iron to the face of another piece of iron, as shown in Figure 3-11, they tend to contract, owing to the tension of the lines. Upon contraction of the lines of force the pieces of iron are pulled toward one another. This is known as the force of attraction and is transmitted by the line to the two pieces of iron, tending to bring them together.

c. Lateral repulsion of lines: Magnetic lines of force repel one another laterally. This phenomenon is also

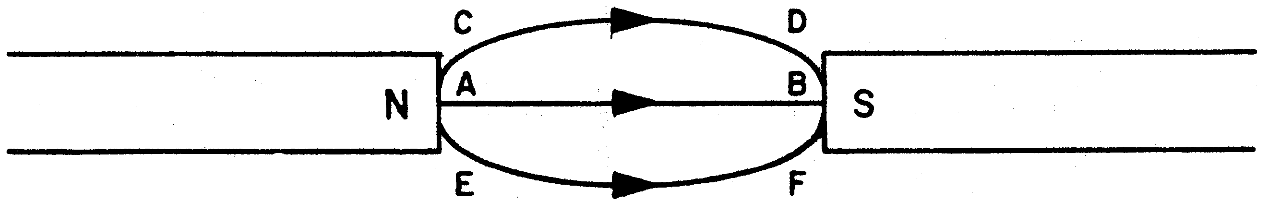


Figure 3-11

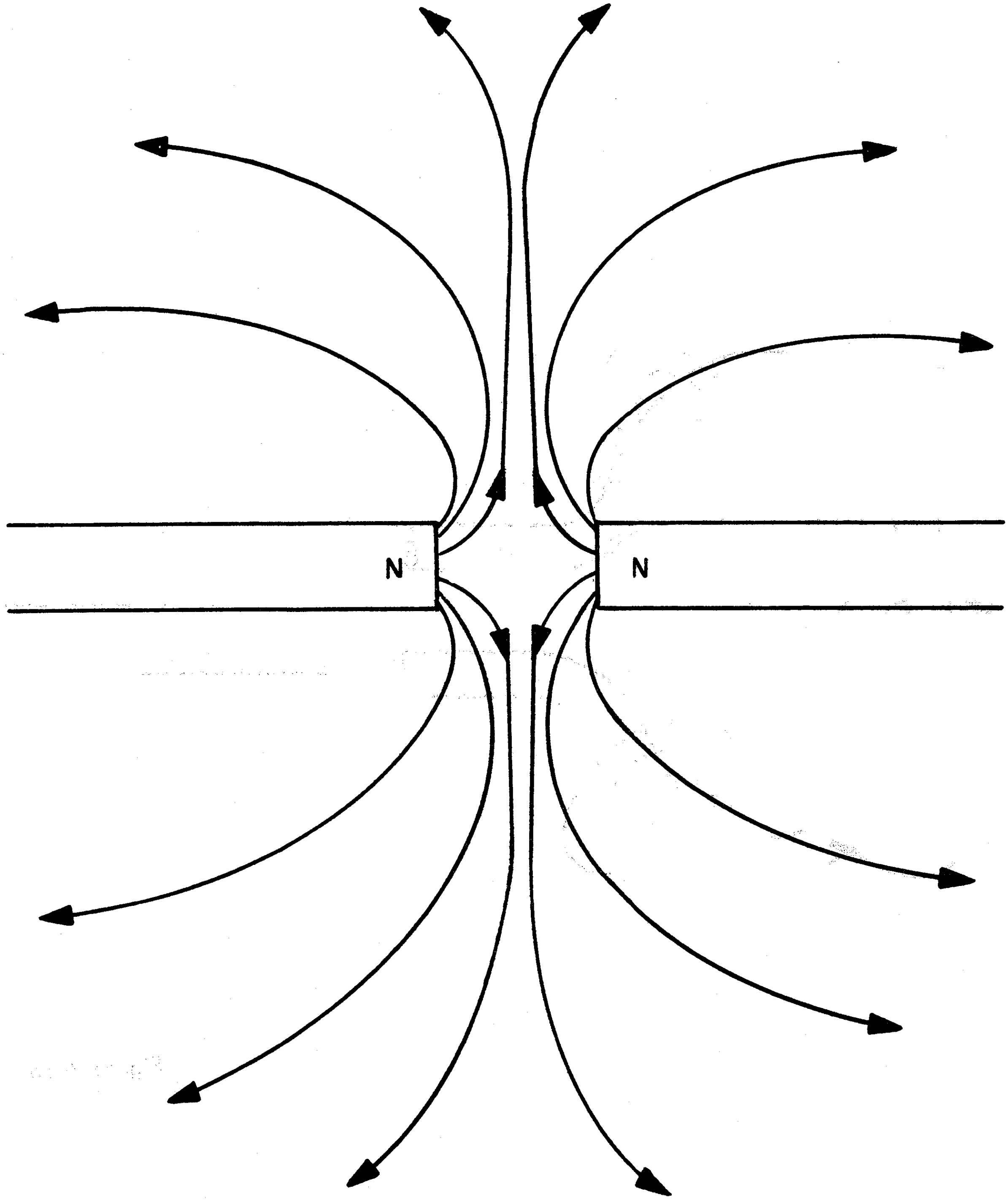


Figure 3-12

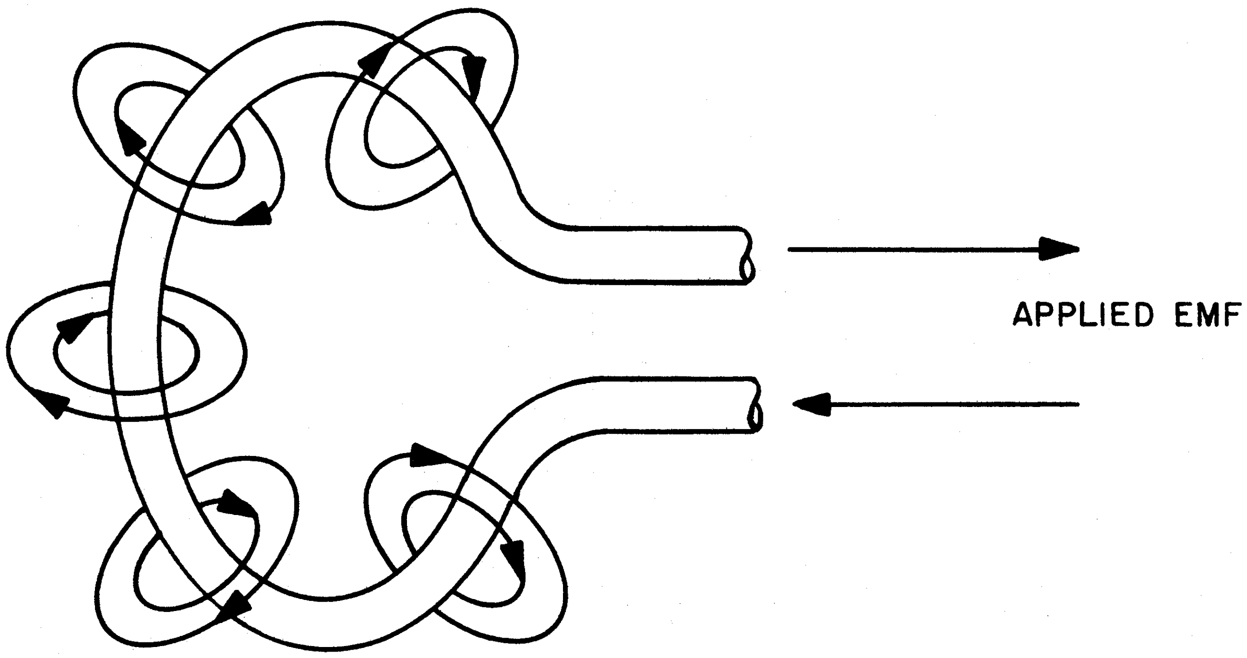


Figure 3-13

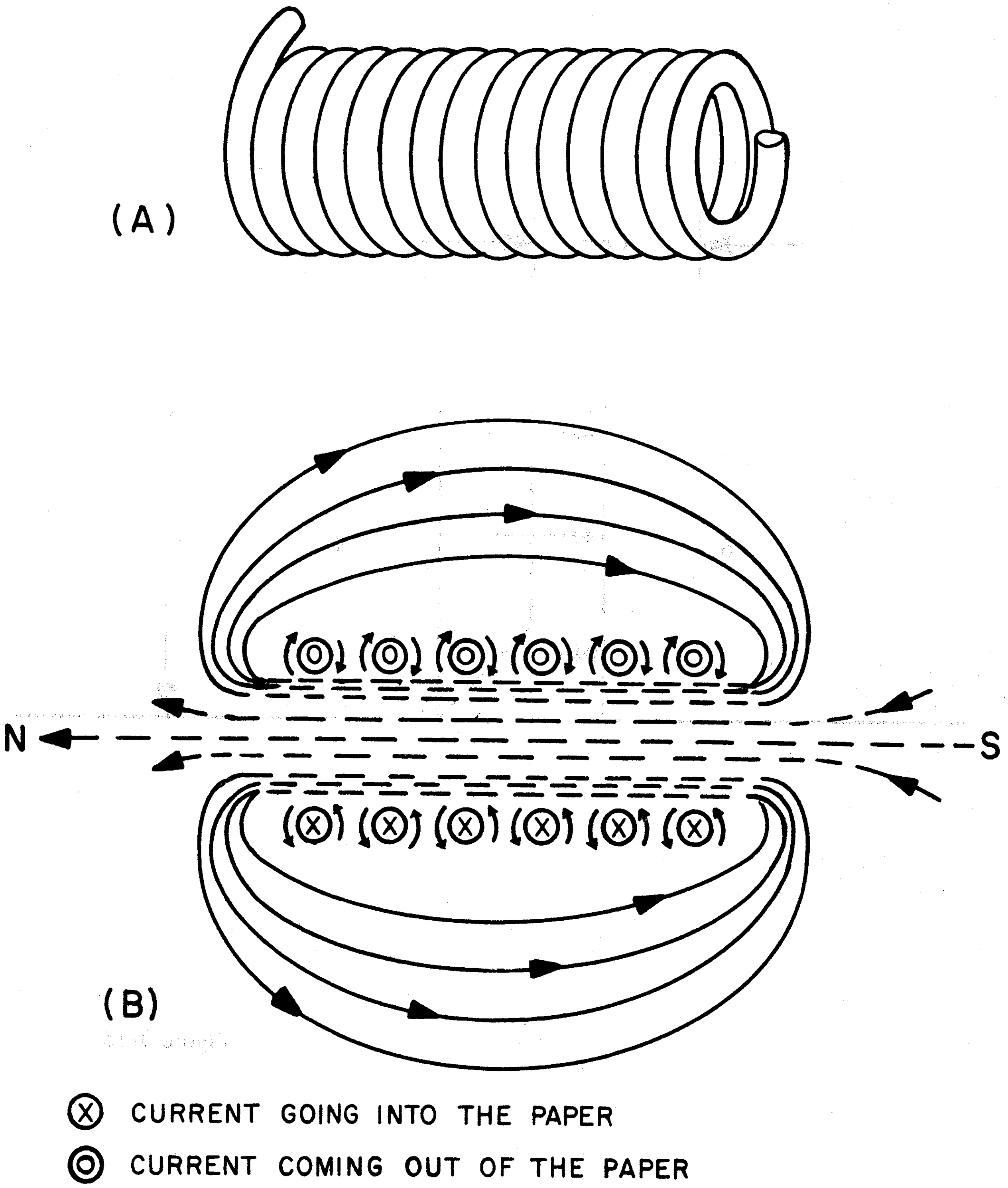

(A)
N
S
(B)
ⓧ CURRENT GOING INTO THE PAPER
◎ CURRENT COMING OUT OF THE PAPER

Figure 3-14

demonstrated in Figure 3-11 where it can easily be seen that line AB repels lines CD and EF laterally, causing them to bow out as shown. In Figure 3-12 it is further demonstrated that if two pieces of iron, both of the same polarity (two south poles or two north poles) are placed opposite each other, then the lines of force repel one another sideways and try to force each other away. This is known as the force of repulsion and is transmitted by the lines to the two pieces of iron, tending to force them apart.

From the results of b and c above it can readily be seen that a law may be stated for magnetic poles similar to that for electric charges. This law is known as the Law of Magnetism, which states: Like poles repel, unlike poles attract.

2.6 SOLENOIDS

In Figure 3-8 the whirls of flux around a straight wire were shown. Suppose this wire is bent into a circular loop across which a voltage is applied as in Figure 3-13. The flux lines still encircle the wire forming the pattern shown. If the wire is further coiled and the turns placed close together as in Figure 3-14(A), the individual magnetic fields that surround each loop will add and the effects will be greatly intensified. This device is known as an electromagnet or solenoid, and is particularly effective if a round bar of soft iron is introduced into the coil as this permits more flux to be set up for a given current and number of turns.

The polarity of a solenoid at any instant may be determined by the left hand rule. Grasping the solenoid with the left hand and pointing the fingers in the direction of the elctron flow through the coils of the solenoid the thumb points towards the north pole of the solenoid. The relationship between current direction and magnetic polarity are shown in Figure 3-14.(B)

Solenoids may consist of a few or many turns arranged in many ways around different types of cores. The type of core to be used is usually determined by the function of the solenoid. When solenoids are used as r-f coils they may utilize air, wooden or plastic cores as coil forms. However, when they are to be used in the capacity of audio and power transformers and chokes, loudspeakers, relays, motors, generators, etc. they use an iron core, which has a high magnetic permeability (see section 2.8.6).

2.7 THE MAGNETIC CIRCUIT

A magnetic circuit is comparable to an electric circuit in almost every respect. As electric current is caused to flow in an electric circuit so are magnetic lines of force (flux) caused to flow in a magnetic circuit. However, the magnetic circuit differs from the electric circuit in three ways, which makes it difficult to obtain the same degree of accuracy when making magnetic calculations as is obtained when making electrical calculations.

In an electric circuit an electric current is confined to a definite path, and since the surrounding air and the insulating supports for the conductors have a very high resistance only a negligible amount of current can possibly be lost. However, since there are no known insulators for magnetic flux it is almost impossible to restrict the magnetic lines of force to definite paths, such as the electric currents are restricted. For instance air which is a good insulator, as far as electric current is concerned is, on the other hand, a good magnetic conductor.

The resistance in an electric circuit is usually constant except under conditions where temperature changes may cause a slight variation. However, when such changes are anticipated the proper compensation can be made. On the other hand, the magnetic resistance of materials is never constant and usually varies over wide ranges.

Large errors may be encountered in magnetic calculations due to the short magnetic paths and their large cross-section in proportion to their length. Usually the geometry of the magnetic paths are so complicated that their magnetic resistance can only be approximated. Therefore, only approximations can be made as to the actual distribution of the magnetic flux.

Although it has been shown above that magnetic calculations can not be made with the order of accuracy that can be obtained with electric circuits, it is still possible to make computations that are close enough for all practical purposes. A thorough understanding of the magnetic units that are associated with magnetic circuits is a prerequisite for obtaining good magnetic calculations.

2.8 MAGNETIC UNITS

The relationship between the magnetic units of a magnetic circuit and the electrical units of an electric circuit are identical to each other. Therefore, it is possible to establish a law for magnetic circuits that is identical in form to Ohm's Law and may be stated as follows: the flux (current) for any given magnetic circuit is directly proportional to the magnetomotive force (emf) of the magnet and inversely proportional to the reluctance (resistance) of the closed circuit. This law is expressed symbolically as:

$$\text{Magnetic Flux } (\phi) = \frac{\text{Magnetomotive Force (mmf)}}{\text{Reluctance (R)}}$$

2.8.1 Magnetomotive Force (mmf)

The magnetomotive force is a force required by a magnetic circuit in order to drive flux (lines of force) through the circuit. This force has its counterpart in an electric circuit (i.e. electromotive force). The magnetomotive force is directly proportional to the product of the current strength and the number of turns in a solenoid. This product is known as the ampere-turns and is abbreviated as NI, although the symbol A.T. is also used. The magnetomotive force only differs from the numerical value of the ampere-turns by a constant factor 0.4π which is approximately equal to 1.257. Therefore, the magnetomotive force is directly proportional to this constant factor and the ampere-turns, and is represented symbolically as

$$\text{Magnetomotive Force} = 1.257NI$$

For example, if 2 amperes are passed through an electromagnet having 200 turns, then the ampere-turns are 2 times

200 which is equal to 400 NI. Thus, the magnetomotive force in the above electromagnet is

$$400 \times 1.257 \times = 503 \text{ gilberts}$$

where a gilbert is the unit of the magnetomotive force.

2.8.2 Reluctance ($\mathcal{R}$)

Reluctance in magnetic circuits is analogous to resistance in electrical circuits. The value of reluctance depends upon the dimensions and the material of the magnetic circuit. No name has been assigned to the unit of reluctance, which is the reluctance (resistance) offered by a portion of a magnetic circuit 1 centimeter long, 1 square centimeter in cross-section, and of unit permeability. The reluctance of a magnetic circuit is directly proportional to its length, inversely proportional to its cross-section, and inversely proportional to the permeability of the material. This is represented symbolically as

$$\mathcal{R} = \frac{L}{A\mu}$$

where L is the length of the path in centimeters, A is the area in square centimeters, and μ (mu) is the relative permeability of the material of the magnetic circuit.

The reluctance of a magnetic circuit differs from the resistance of a electric circuit in one respect. The resistance of an electric circuit does not depend on the magnitude of flow, except as it is influenced by heating, whereas reluctance depends on the magnetic flux density (see section 2.8.4).

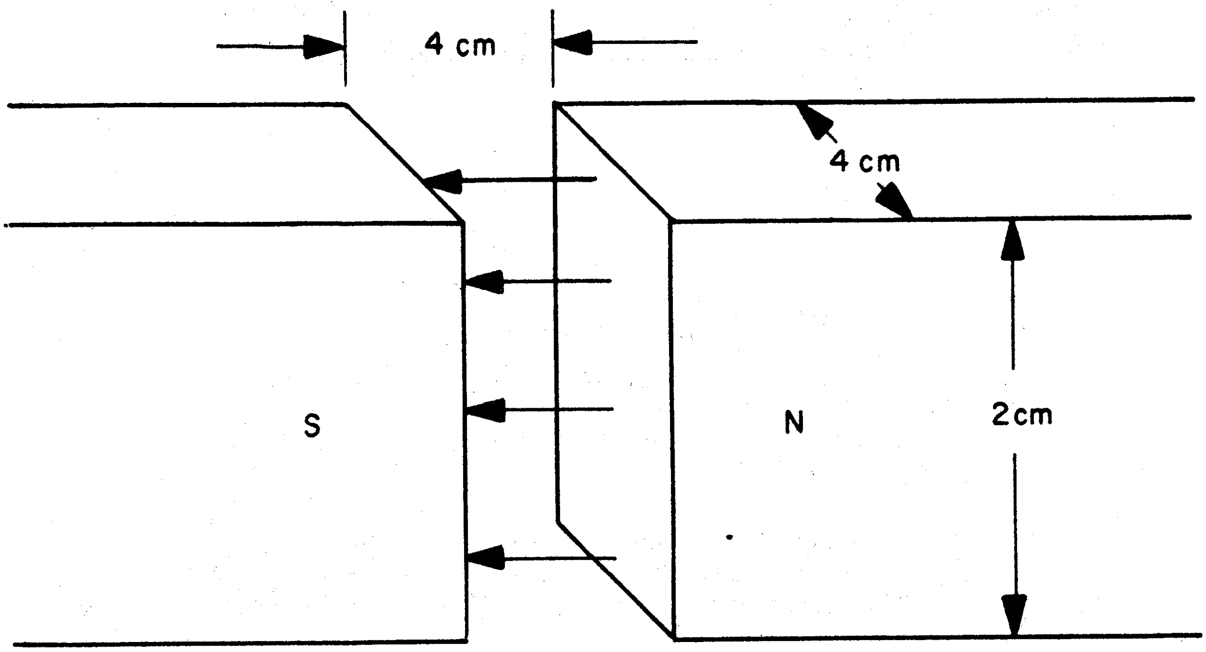


Figure 3-15

When the magnetic flux density(B), of a magnetic circuit is altered the permeability of the material is changed, which in turn either raises or lowers the reluctance.

The following example demonstrates how reluctance is determined. By considering an air space (permeability equal to 1) between two magnetic poles Figure 3-15, having a length of 4 centimeters and a cross-section of 2 centimeters by 4 centimeters or 8 square centimeters, the reluctance is:

$$\text{Reluctance} = \frac{4}{1 \times 8} = \frac{1}{2} \text{ unit.}$$

2.8.3 Flux (ϕ)

Any magnetic circuit has a magnetic field surrounding it, which is made up of magnetic lines of force. These lines of force that cross a given air space or field are known as the magnetic flux ϕ, (phi). The unit of magnetic flux is the maxwell. One maxwell is one line of magnetic flux. The magnitude of a line of flux may be expressed in several ways in terms of other phenomena such as its action on a permanent magnet of known strength, or by its effect on a current-carrying conductor of given length and having a given current flowing through it.

It is convenient at this time to define the maxwell in terms of magnetomotive force and reluctance by means of a law, similar to Ohm's Law for electric circuits, which was discussed in section 2.8. Thus, if a magnetomotive force of one gilbert is impressed on a path having one unit of reluctance, then one line or maxwell is set up in this path.

2.8.4 Flux Density (B)

Flux density B, is the amount of flux per unit area. The unit of flux density is the gauss which is equal to one maxwell per square centimeter. If an example is taken where the flux in a given air gap is 10,000 maxwells and the cross-sectional area of the gap is 75.0 square centimeters, the flux density is:

$$B = \frac{10{,}000}{75} = 133.3 \text{ gauss}$$

In a substance such as iron the reluctance is variable and depends upon the flux density. When the flux density is too high, the iron becomes saturated (any further increase in the magnetization force has little or no effect on the orientation of the molecular magnets in the iron) causing the permeability to decrease and the reluctance to increase. Therefore, it is necessary, when using iron cores, for the designer to proportion the area of the iron core so as to prevent too high a value of flux density. Important consideration must always be given to the amount of flux passing through each square centimeter of the material (flux density), in the magnetic circuit rather than the amount of flux passing through it (magnetic flux).

2.8.5 Permeance ($\mathcal{P}$)

Permeance is the property of a magnetic circuit which permits the passage of magnetic flux. It is the reciprocal of reluctance and is therefore equal to $1/\mathcal{R}$.

It is analogous to conductance in an electric circuit.

2.8.6 Permeability

The permeability μ (mu), of a material, is the ratio of the flux that exists in the material to the flux that would exist in the same space if the material was replaced by a vacuum. A vacuum has a permeability of one (unity) and with the exception of iron, steel, nickel, and certain iron oxides most of the other materials have a permeability of one. Commercial iron and steel have permeabilities that range from 50 and even lower to about 2,000.

2.8.7 Field Intensity (H)

The field intensity H (the magnetic field intensity in air) is directly proportional to the current flowing in the winding of an electromagnet. If the electromagnet is very long as compared to its diameter, and is constructed with one turn of wire to each centimeter of length, and the current in the winding is one ampere, the field intensity in the air on the inside of the electromagnet can be shown to be equal to $.4\pi$ or 1.257 gilberts.

A gilbert, as has already been established, is the unit of magnetomotive force. Therefore, each turn of wire per centimeter for a current of one ampere gives 1.257 gilberts, and the intensity H increases proportionally to the current and to the number of turns of wire per centimeter of length. Thus, the field intensity is equal to the magnetomotive force per centimeter of circuit which is represented symbolically as

$$\text{Field Intensity (H)} = \frac{\text{Magnetomotive Force (mmf)}}{\text{Length of Electromagnet in Centimeters}}$$

The unit of field intensity is one gilbert per centimeter and is called the oersted.

2.9 B-H CURVE

Mention was made in section 2.8.4 that the reluctance of iron and other ferromagnetic materials vary with the flux density. Therefore, the flux density is not in direct proportion to the magnetomotive force, as is the case of an air path. Thus, it is necessary to determine the flux density for several different values of impressed magnetomotive force, and plot a curve between the two. This curve may then be used to determine the magnetic properties of the material at any flux density.

Since it has already been established that the amount of flux depends upon the magnetomotive force and the length and cross-section of the iron under test, it is necessary to express the relationship between the magnetomotive force and flux on a per unit basis, i.e., for one square centimeter of cross-section and one centimeter of length. However, since the flux per square centimeter is the flux density B, and magnetomotive force per centimeter is the magnetizing force H, the curve is known as the B-H curve.

The magnetization curve that is obtained when the B and H values are plotted depends upon many conditions such as: (a) the type of metal used as the sample; (b) the degree of purity of the metal; (c) the heat treatment

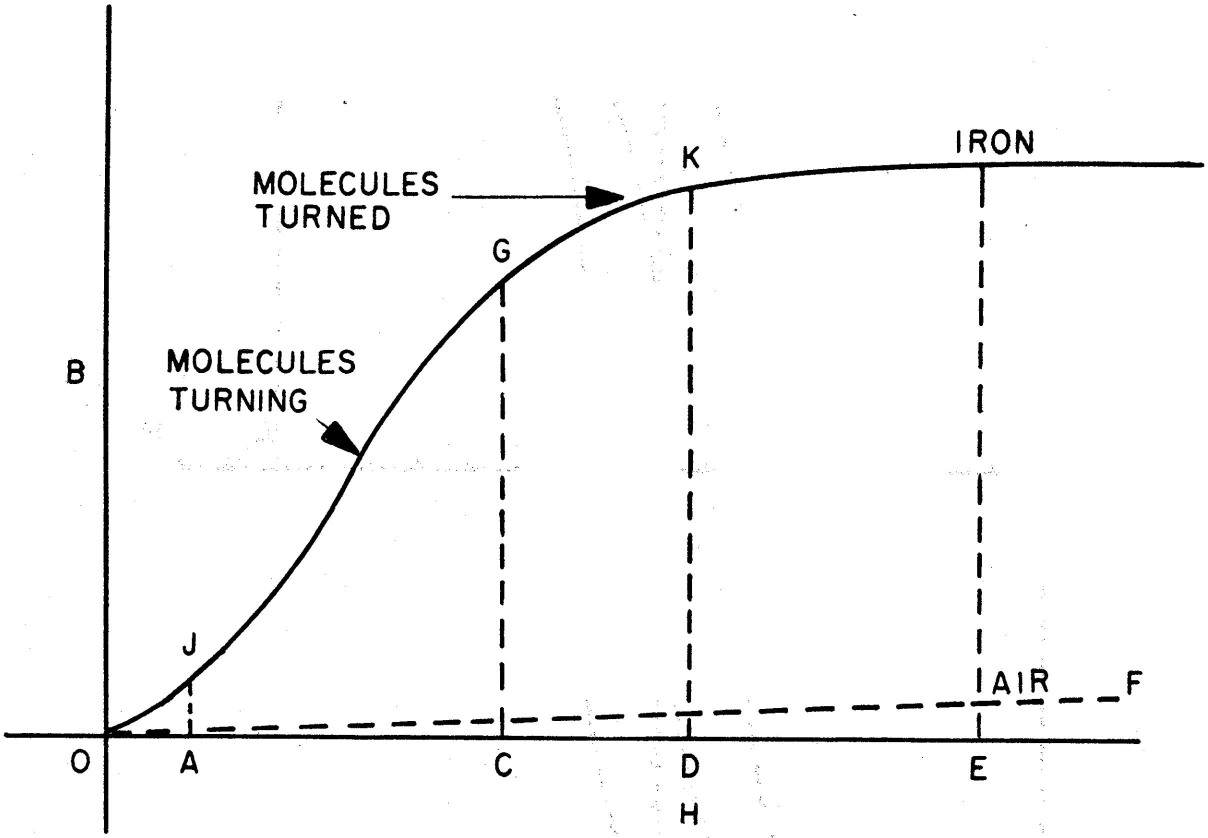


Figure 3-16

used in preparing the metal; (d) the previous magnetic history of the metal. That is, whether or not it has ever been subjected to a high degree of magnetization.

Plotting B versus H for increasing values of H from zero to E, as shown in Figure 3-16, gives rise to a curve known as the normal curve. B increases very slowly for the small values of H, (from 0 to A) but then increases rapidly as H increases from A to C. However, B increases again less rapidly as H increases from C to D, which is due to the saturation of the metal taking place. This portion of the curve is known as the knee of the characteristic.

With further increase of H from D to E, B increases very slowly, just as though it had the same reluctance as air from that point on. Since it can not carry flux on the large scale it did before, the iron is said to be saturated. However, it is obvious from the graph that it can still function as an equal volume of air would. The B-H curve for air is represented in Figure 3-16 as the broken line OF. This is, for all practical purposes, a straight line because the reluctance of air is constant. Therefore, there are no saturation effects and B is directly proportional to H.

2.10 HYSTERESIS

When a magnetomotive force or magnetizing force was shown to act on an iron sample from zero to E in Figure 3-16 a normal saturation or magnetization curve was obtained. However, if the magnetization force is now decreased (see Figure 3-17), the magnetization curve OC does not decrease along the line of ascent, but decreases less rapidly along CD. Therefore, when H is reduced from OA to zero, there is still some flux set up in the material of a value of OD. This is called the residual flux density or remanent flux density, which causes the iron sample to act like a permanent magnet. Before this remanent flux density can be decreased to zero, the magnetizing force must be reversed in direction. That is, a negative magnetizing force OE is required in order to reduce the remanent flux density to zero. This value of H, namely CE, is called the coercive force as it represents the counter or negative force required to reduce the maximum remanent flux density to zero.

If the magnetizing force is now increased in the negative direction to F (where OF = -OA), the flux density increases to a negative maximum of FG. This negative maximum is equal and opposite to the positive maximum flux density AC. If the magnetizing force H is varied from a value of OF toward a value of O, then the flux density decreases as shown by the curve, reaching the value of OI as the magnetizing force reaches O. Notice that OI = -OD. As the magnetizing force is increased toward a value of OA, the flux density falls to O and then rises in the opposite direction, reaching the value AC as the magnetizing force reaches the value OA. This lag of B with respect to H is given the name hysteresis and the closed curve or loop that is formed is known as the hysteresis loop. This loop is symmetrical if the iron sample is carried through the same cycle of magnetization several times.

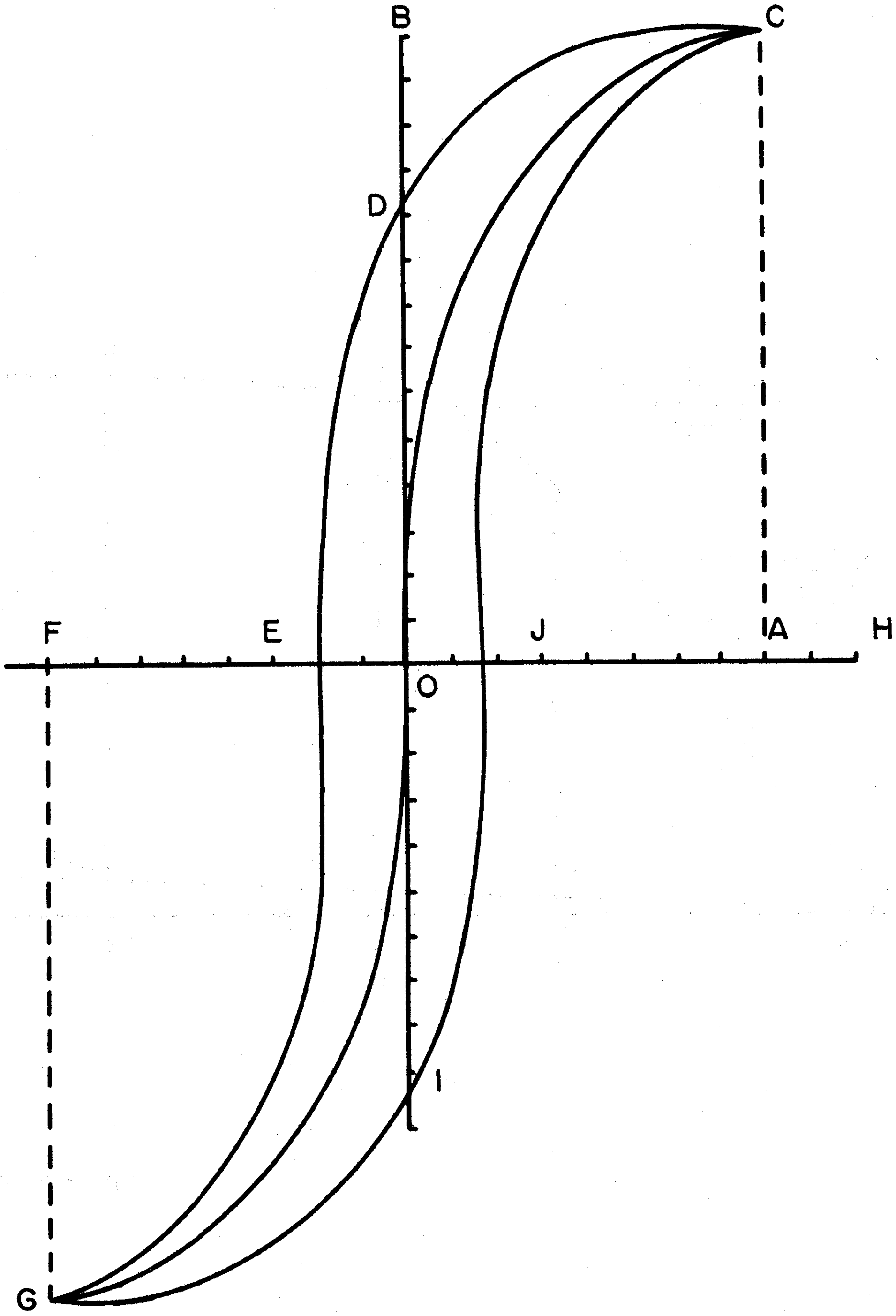


Figure 3-17

Different types of materials give rise to different shaped hysteresis loops when under the influence of some magnetizing force. The loop shown in Figure 3-17 is the type of curve obtained from iron; however, in computer storage systems where a high residual magnetism is desired it is most desirable to use materials that give a relatively more square hysteresis loop.

Different materials follow certain magnetic curve paths under the influence of a magnetizing force. The reason for this variety in magnetization curves can be explained with the use of the Weber and Ewing theory which has been discussed in section 2.3. To further elaborate on this theory, it is obvious that as an outside magnetic field is applied to a specimen of magnetic material the individual molecules align themselves more and more in the direction of the applied field, with each individual magnetic field adding up to the total magnetic field of the specimen. Each degree of magnetization can be followed by inspecting Figure 3-16 which shows the action taken by the molecules from the unorientated state to the state of complete magnetization. It is the amount of alignment of the molecules that still exists after the magnetization force is reduced to zero that determines the residual magnetism retained by the material under test. It is the extent of residual magnetism that determins how and where the material can be utilized. For instance, as has already been mentioned, it is the materials that have a square hysteresis curve that are useful in computer storage systems.

2.10.1 Hysteresis Loss

Since hysteresis is the tendency for a material to stay magnetized after a magnetomotive force has been applied and removed, the loss in energy which must be applied to overcome this residual magnetism is called the hysteresis loss. This loss which is in the form of heat generated in the magnetic material is proportional to the area of the hysteresis loop, and occurs only when the magnetizing force H is alternating in character. Therefore, hysteresis is an important characteristic to be condisdered in transformers, A-C iron core chokes, etc. Although it is possible to calculate this loss by finding the area of the loop to scale and dividing by 411, it is generally furnished by the manufacturer for the particular material offered, and at a certain flux density. There is a smaller hysteresis loss in soft iron than in the other forms of iron because soft iron offers less opposition to changing magnetism than the harder forms of iron.

2.10.2 Eddy Current Losses

There exists not only a hysteresis loss in cores subjected to a varying or an alternating magnetic flux, but eddy current losses as well. Eddy current losses in magnetic materials are due to the voltages and consequent circulatory currents induced by the variation of the flux. Eddy currents are the I^2R losses that occur in the material. These currents are very large, because the

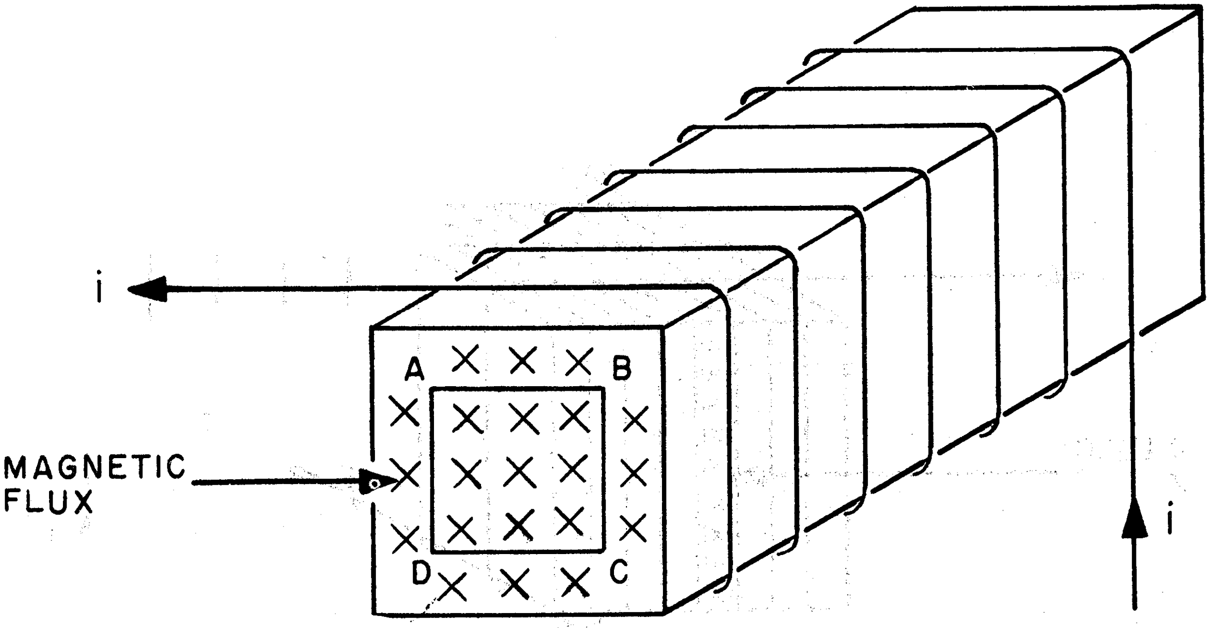


Figure 3-18

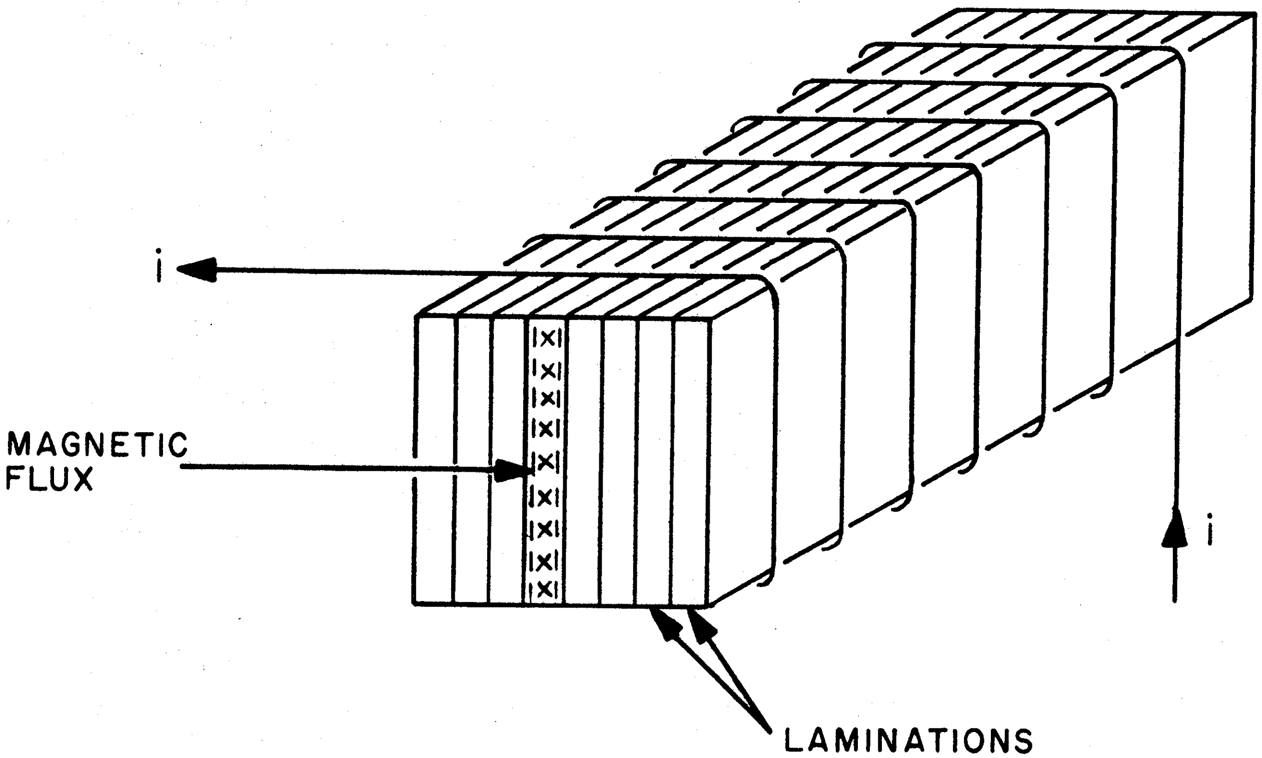


Figure 3-19

resistance of the path in the magnetic material is very low due to the large solid cross-sectional area, even though the induced voltages are very low.

Eddy currents may occur in any conducting substance. However, in order to demonstrate how eddy currents are set up in iron or other material used for magnetic circuits when they are subjected to cyclic magnetization, a piece of iron is shown in Figure 3-18. An electron flow in the coil causes the magnetization of the iron bar and the marks indicate the magnetic flux going into the iron bar. This flux increases and decreases with the electron flow and reverses in direction when the electron flow reverses. Considering a path a-b-c-d in the plane of the cross-section any increase of the flux in the direction shown induces in this path (and in any other path linked by the changing flux) an emf in such a direction as to cause a counterclockwise electron flow around the path. Due to this electron flow the resistance of the path converts the energy into heat. Therefore, by summing up the I^2R losses in all such elemental paths as a-b-c-d, commencing with a very small path near the center of the cross-section and ending with a path lying just under the surface the total I^2R loss represents the eddy current loss.

It has been determined, however, that this eddy current loss may be sufficiently reduced by building up the required cross-section for the flux path by stacking thin pieces known as laminations (see Figure 3-19). Due to the small magnitude that the conductor cuts across magnetic lines of force, sets up an electric field in the conductor.

Relative motion between a magnetic field and a conductor can be produced mechanically either by moving a of the emf's set up in the material by varying the flux, the natural oxide on the surface of the sheet iron or steel from which the laminations are punched effectively insulate the laminations from one another. Therefore, each eddy current path is limited to a single lamination.

2.10.3 Core Losses

The core losses that exist in any core which is subjected to a varying or an alternating magnetic flux are made up of the hysteresis and eddy current losses. As it has already been pointed out both hysteresis and eddy current losses can be sufficiently reduced by the proper selection of material and by laminating.

Since the rating of transformers (see Section 2.12 of this Part) is based, to a considerable extent, on the core losses, a low core loss is essential. Therefore, silicon steel because of its low core loss is almost used entirely in transformers. When approximately 4 per cent of silicon is alloyed with steel the hysteresis loss is reduced materially, and since silicon also increases the electrical resistivity, the eddy current loss is reduced.

2.11 ELECTRO-MAGNETIC INDUCTION

2.11.1 Generation of EMF

As noted in Section 2.4 of this Part, a moving electric field creates a magnetic field. It can also be shown that motion of a magnetic field with respect to a conductor, such

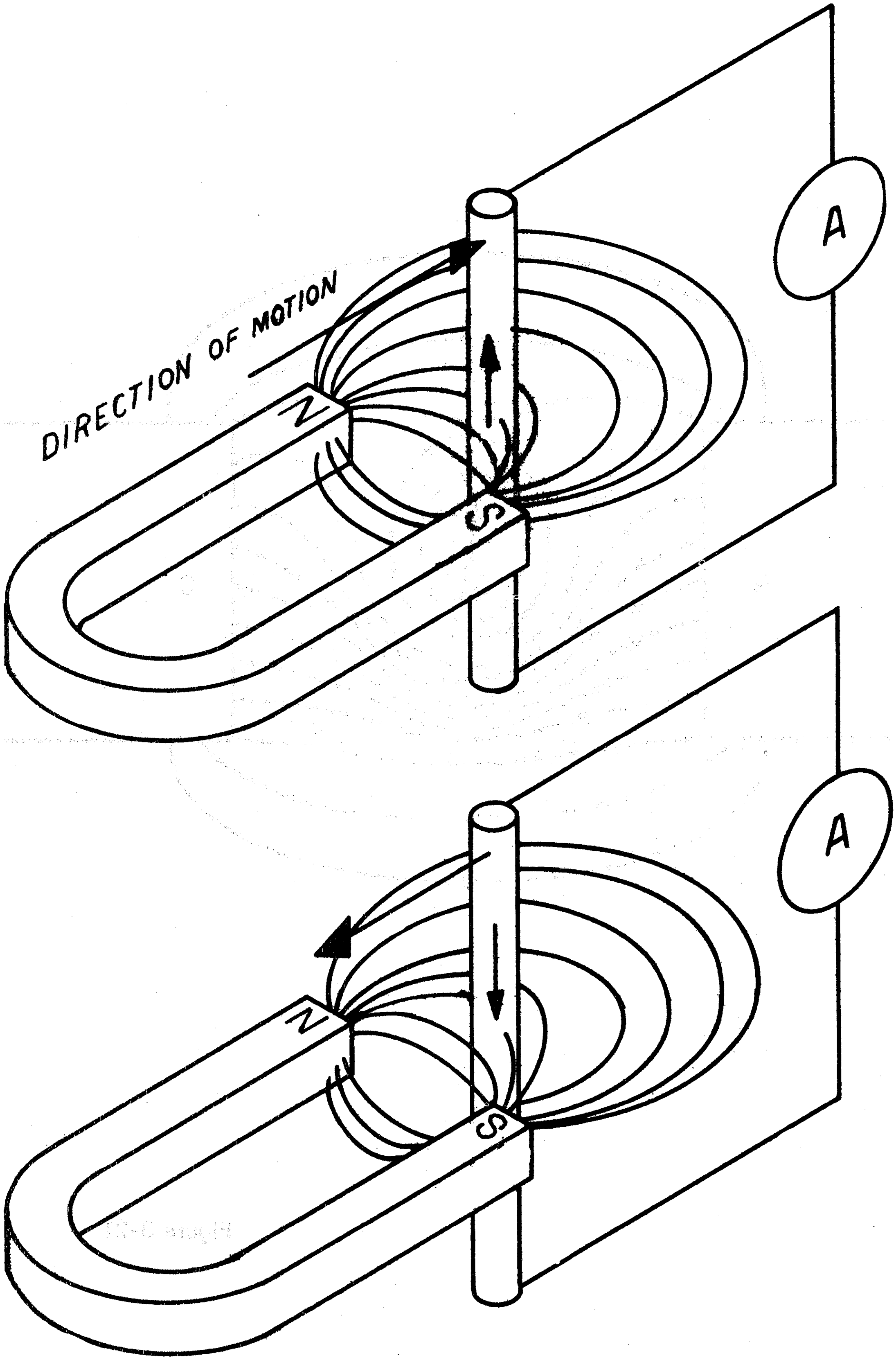


Figure 3-20

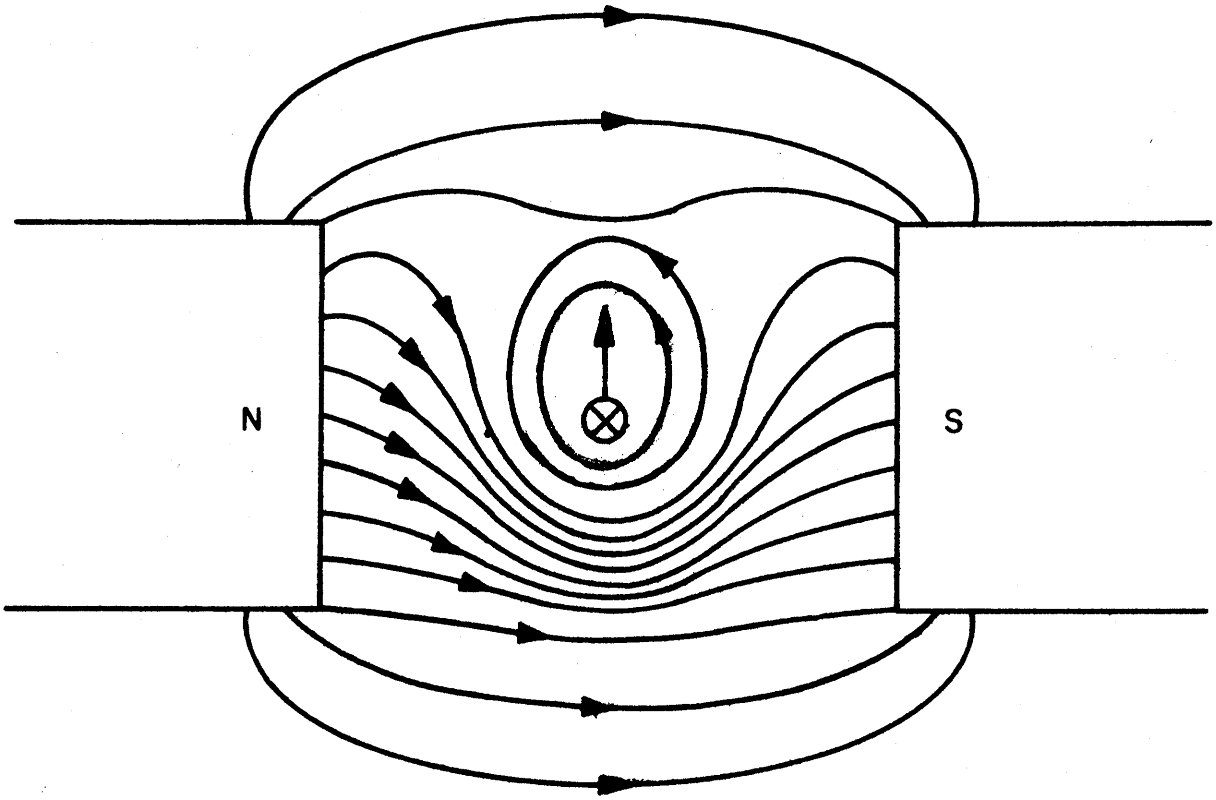


Figure 3-21

magnet past a stationary conductor or by moving a conductor past a stationary magnet, or it can be produced electrically by increasing or decreasing the flow of current through an electro-magnet.

Mechanically produced electro-magnetic induction is the underlying principle of the generator which supplies most of the electric power in the world today.

Electrically produced electro-magnetic induction is the underlying principle of the transformer which is one of the most important of circuit components.

The force that is produced in a conductor by electro-magnetic induction is called the induced emf (or induced voltage). If the conductor is part of a closed circuit, then the induced emf causes a current to flow through that circuit.

2.11.2 Lenz's Law

The polarity of the induced emf is always such that the resultant current creates a magnetic field which opposes the motion of the field inducing the emf. This is known as Lenz's Law.

The implications of Lenz's Law with respect to mechanically produced electro-magnetic induction are illustrated in Figure 3-20. Here a conductor is assumed to be moved through the field of a permanent magnet as is shown. The induced emf is such as to produce electron flow in the directions indicated by the arrows. Figure 3-21 shows how the field created by the current through the conductor and the field of the permanent magnet interact to produce a mechanical force which opposes the downward movement of the conductor. If the conductor is moved upward through the field, the direction of the electron flow through the conductor is reversed so that the interaction of the two fields is such as to oppose the upward motion.

The relationships described above can be remembered in terms of the following rule of thumb:

Extend the thumb, forefinger and second finger of the left hand at right angles to each other. Point the thumb in the direction of the magnetic lines of force. Point the second finger in the direction of motion of the conductor. Then the forefinger points in the direction of electron flow through the conductor.

2.11.3 Self-Induction

When the current through a conductor increases, lines of flux move outward from the center of the conductor in the same manner that ripples move outward from a spot where a stone is dropped into water. These lines of flux cut across the conductor itself inducing an emf in the conductor which opposes the increase of current flow. Similarly, when the current through a conductor decreases, the lines of force collapse toward the center of the conductor cutting across the conductor itself and inducing an emf which opposes the decrease of current flow. The opposition offered in each case by the so-called counter emf of self-induction to the change in current which causes it, corresponds to the statement of Lenz's Law in the previous section.

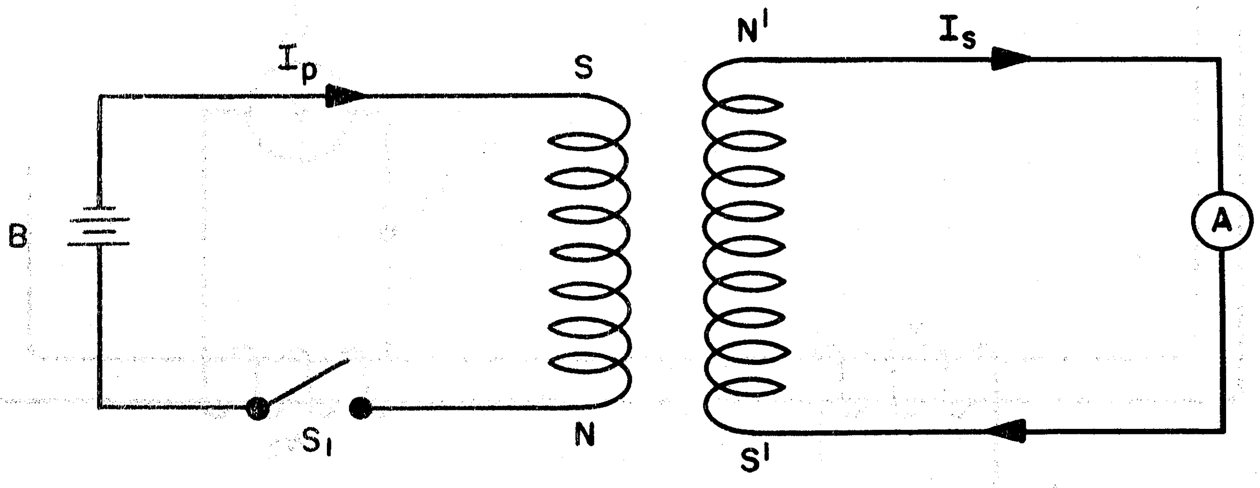


Figure 3-22

When a straight conductor is coiled into the form of a solenoid (see Figure 3-14) then the expanding or contracting lines of force associated with each turn cut across adjacent turns, greatly increasing the magnitude of the counter emf of self-induction. Thus, a solenoid offers much more impedance to changes in current than does a straight conductor. This impedance is of great importance in alternating current circuits as explained in the next chapter.

2.11.4 Mutual Induction

If a magnetic field of one circuit cuts across a conductor in a second circuit, it induces an emf in that second circuit. The current which flows in the second circuit, in response to the induced emf, sets up a magnetic field which opposes the field of the first circuit. This so-called mutual induction which is the basis of transformer action is shown in Figure 3-22.

If switch S1 is closed, then battery B forces current through the associated circuit creating a strong expanding magnetic field. This expanding field cuts across the turns of the winding belonging to the second circuit, which also includes ammeter A, causing current flow through that circuit in the direction shown. The polarity of the field generated by the battery or primary circuit is indicated by the notations N and S. The polarity of the opposing field set up by the current through the secondary circuit is indicated by the notations N' and S'. Notice that the fields oppose each other.

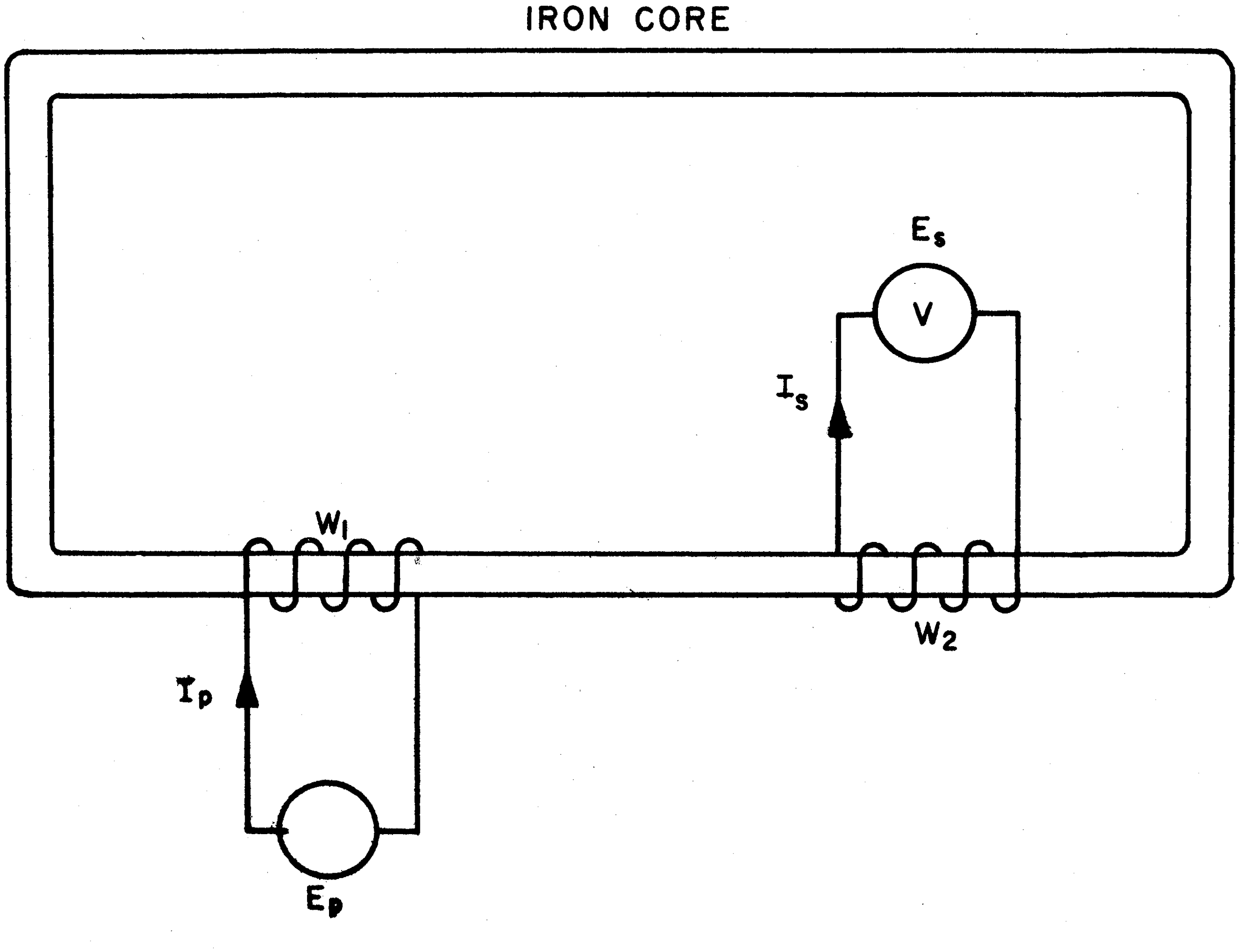


Figure 3-23

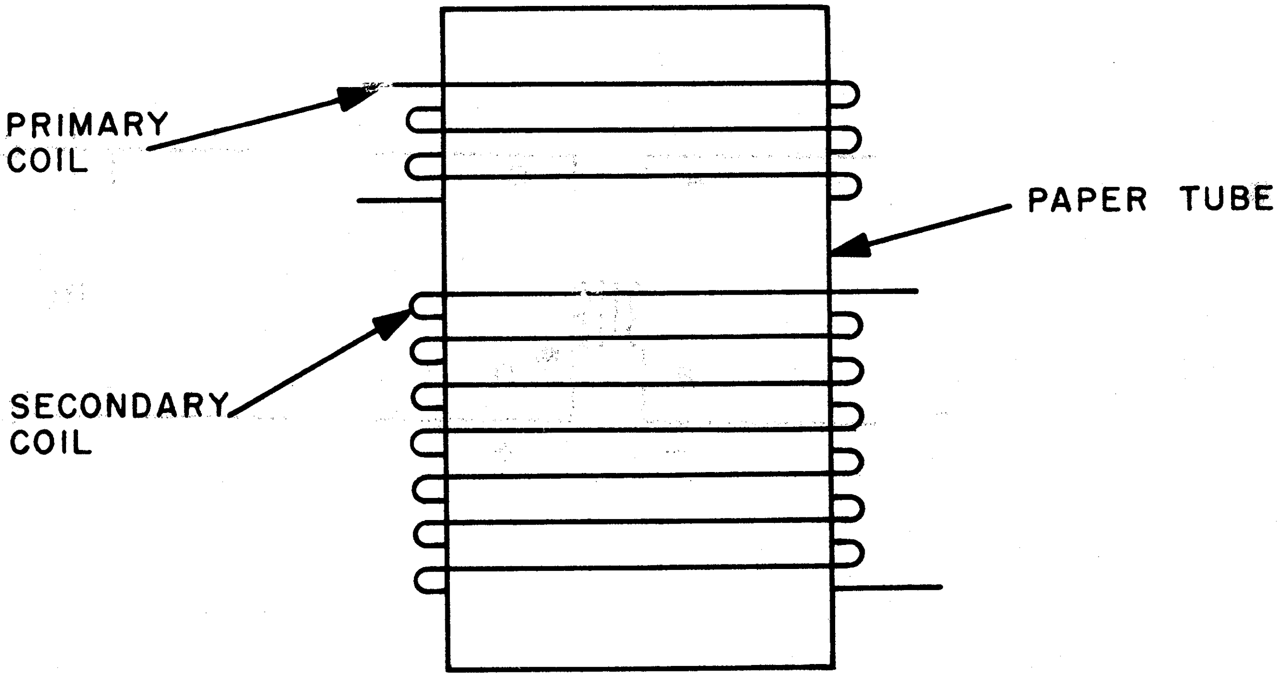


Figure 3-24

2.12 TRANSFORMERS

As it has already been pointed out in section 2.11.2, when a magnetic field of one circuit cuts across a conductor of another circuit an emf is induced in the second circuit. This was called mutual induction which is the basis of transformer action.

A transformer usually consists of two or more coils so placed that mutual inductance exists between them. The coils may be either wound upon a ferromagnetic core which is part of a magnetic circuit, Figure 3-23, or they may be wound upon a non-magnetic form, such as a paper tube, Figure 3-24. Transformers which are used at power frequencies and audio frequencies are built with ferromagnetic cores and the windings are usually wound so that the greatest percentage of the windings are linked by all the flux.

If two coils of wire are arranged on a common iron core, Figure 3-23, so that they are close together and a sine wave alternating voltage Ep is impressed on coil W_1 a corresponding a-c electron flow is caused in W_1. This electron flow is of such a magnitude that it sets up an a-c flux which in turn induces a counter voltage in W_1 that just balances Ep. This a-c flux is also assumed to pass through coil W_2 in traversing the high permeability iron core, inducing a voltage. This voltage induced in coil W_2 is called the secondary voltage and the coil W_2 the secondary coil, whereas coil W_1 is called the primary coil. This arrangement, including the iron core, is known as a transformer, and the two coils are said to be magnetically coupled.

2.12.1 Transformer Principles

If in Figure 3-23 windings W_1 and W_2 have the same number of turns and it is assumed that the coils and the core are so constructed as to have negligible energy losses, a voltmeter would give the same reading across each coil. Thus, the emf of the secondary coil is equal to the impressed emf of the primary coil. This is known as an ideal transformer of unity ratio. However, by increasing the number of turns of the secondary coil W_2 the voltmeter reading on the secondary would be greater than the reading on the primary side of the transformer. Now if the situation is reversed (more primary turns than secondary turns) the effect is the reverse. Therefore, the voltage across the two coils is directly proportional to the number of turns. This relation is symbolically represented by the equation

$$\frac{V_p}{V_s} = \frac{N_p}{N_s}$$

where V_p is the voltage across the primary, V_s the voltage across the secondary; N_p is the number of turns comprising the primary; and N_s is the number of turns comprising the secondary.

Suppose 10 volts (E_p) is impressed across the primary coil W_1 in Figure 3-23, which is comprised of 20 turns. Then the voltage that is induced across the secondary coil, which is comprised of 40 turns, is;

$$\frac{10}{V_s} = \frac{20}{40}$$

$$20\ V_s = 400$$

$$V_s = 20 \text{ volts}$$

This clearly illustrates the fundamental transformer action, that two coils coupled together can transform an impressed voltage across one coil into a higher voltage across the other coil. The reverse is also true.

The current in the primary and secondary coils, like the voltage, depends upon the ratio of the number of turns of the primary to the number of turns of the secondary. The relation between the current values is the inverse ratio of the number of turns. Therefore, the winding that has the greater number of turns has a proportionately smaller current. This can be seen by considering the law of conservation of energy, where the energy existing in the secondary circuit can never exceed, but for an ideal transformer just equal the energy of the primary circuit. This can be shown by examining the following equations where since

$$P_p = P_s, \text{ and } P = EI$$

it follows that

$$V_s I_s = V_p I_p$$

from which

$$\frac{I_s}{I_p} = \frac{V_p}{V_s} \quad \text{or} \quad \frac{I_s}{I_p} = \frac{N_p}{N_s}$$

where P_p is the energy existing in the primary coil; P_s is the energy existing in the secondary coil; P is the power; E is the emf; I is the current; I_s is the current in the secondary; and I_p is the current in the primary.

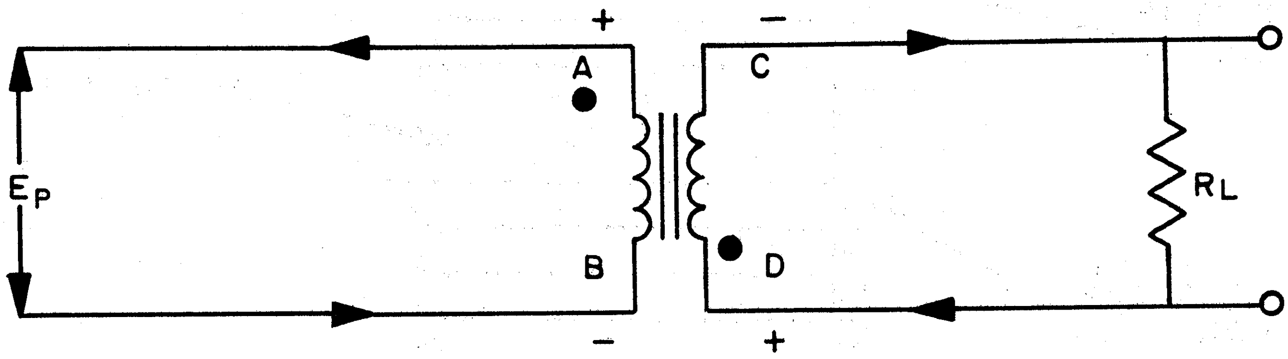


Figure 3-25

2.12.2 Polarity Relations

One important transformer characteristic is the reversal of polarity of the induced secondary voltage. Consider the condition in Figure 3-25 where the applied voltage E_p has an instantaneous polarity such that the top terminal "a" of the primary is positive and the bottom "b" is negative. In all cases the secondary induced voltage is in such a direction that the current it produces in the winding and R_L is in the proper direction to oppose, magnetically, the action of the primary current. The result is that the phase of the secondary current is such that it is opposite at all times to that of the primary current. This introduces the important action called phase inversion. Thus in the case of the example given, the polarity of the secondary current is reversed as shown at "c" and "d".

In an actual transformer, the polarity of the terminals depends not only on the direction of the current as illustrated by the arrows (Figure 3-25) but also on the physical arrangement of the coils on the core form.

Thus if it is desired, a transformer can be made that does not exhibit any reversal of polarity. The dots shown in the figure are one convention used at times to emphasize phase difference between primary and secondary currents.

2.12.3 Iron-Core Transformers

The transformers used in the previous discussion were assumed to possess iron cores. That is, the windings were wound on iron forms. In practice, iron core transformers are used only in applications where they operate at low frequencies; i.e. over a range of from 20 cps or less to 20,000 cps because the iron helps to obtain high inductance and hence high inductive reactance in spite of the low frequencies involved, and without requiring a prohibitive number of turns. A high inductive reactance, in turn, serves to keep the magnetizing current at an acceptably low value relative to the reflected primary load current, so that the unit approaches the ideal transformer in its characteristic.

2.12.4 Air Core Transformers

At the higher portion of the spectrum, such as at frequencies used in radio work (up to and above 5,000,000 cps), iron cores tend to have rather high losses. As a result, the coils are generally wound on hollow dielectric (insulating) forms or tubes. The core is essentially air or its equivalent in magnetic permeability, and the devices are known as air-core transformers.

2.13 TYPES OF TRANSFORMERS

2.13.1 Step-Up and Step-Down Transformers

Assume, as before in Figure 3-23, that W_1 is the primary. Then if W_2 has more turns than W_1, the secondary voltage is greater than the primary voltage, and the device is a step-up transformer.

On the other hand, if W_2 is the primary and W_1 the secondary, the same unit acts as a step-down transformer in that the voltage induced in W_1 is less than the voltage applied to W_2.

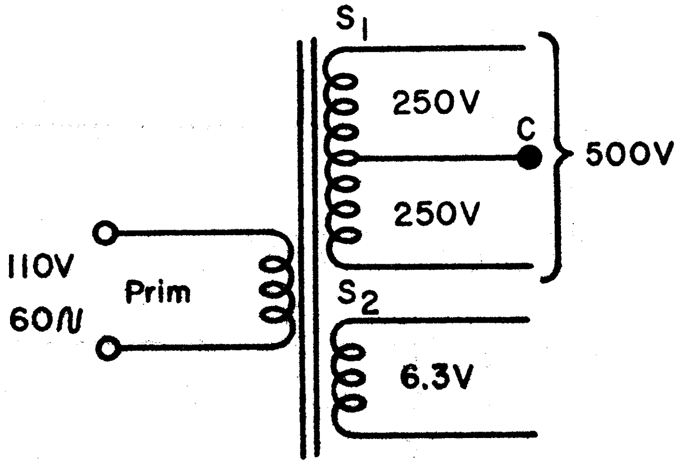


Figure 3-26

In a transformer, more than one secondary can be wound on the same core; the varying flux, in traversing each winding, induces a voltage in the winding that is proportional to the number of turns. Thus, the voltage may be stepped up in one secondary and stepped down in another.

As an example, refer to Figure 3-26 where a power supply transformer is shown. The primary is connected to 110-volt a-c power supply. The secondary S_1 has many more turns so that 500 volts are developed across the full winding. At the midpoint of the winding is soldered a connection C; the voltage from either end to C is 500/2 equal to 250 volts. Tap C is known as a center tap. Winding S_1 furnishes the high voltage which, after rectification, becomes a d-c source.

Secondary S_2 is also a secondary winding furnishing 6.3 volts which is normally used to feed the heaters of vacuum tubes in the system in which it is used.

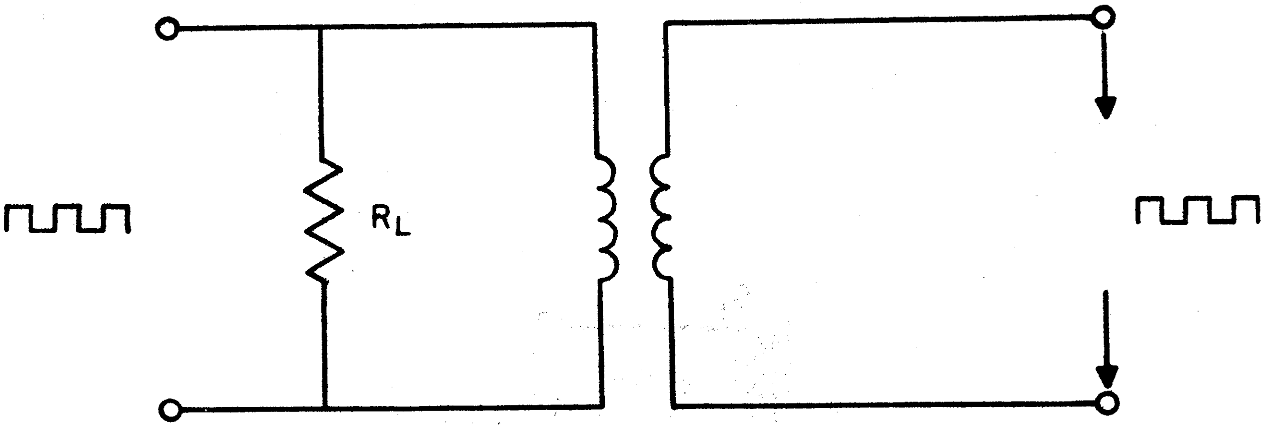
2.13.2 Auto Transformers

The auto transformer is one in which one winding serves as both primary and secondary. The input voltage is applied to the two ends of the coil and the output is taken between one end of the winding and a tap at some intermediate point. Auto transformers are inexpensive and small, but, of course, they do not afford the isolation between the primary and secondary circuits that the ordinary transformer does. Thus, although they afford the advantages of greater efficiency and better voltage regulation, they have the disadvantage of conductive connection between the high voltage and low voltage circuits. One common use of auto transformers is to furnish a reduced voltage for starting electric motors. As components of these so-called compensator starters, they are equipped with switching allowing the lapped position to be varied so as to select different transformation ratios during motor starting.

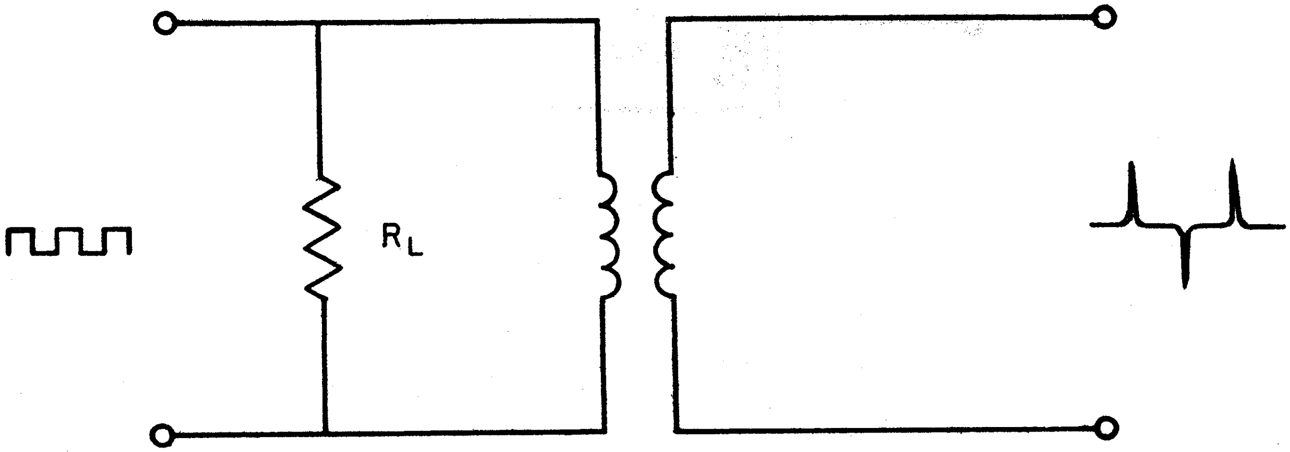
2.13.3 Impedance Matching Transformers

A transformer may be used as an impedance matching device. Practical applications of this type are enormous in number as the following simple example helps to make clear. A vacuum tube power output stage delivers maximum output into a loudspeaker without exceeding the percentage distortion specified, if it feeds its power into a specific value of load resistance R_L.

On the other hand, a loudspeaker may have a resistance quite different from R_L. If the resistance of the loudspeaker is designated R_{LS} it is made to look like R_L by interposing a transformer of the proper turns ratio between the amplifier stage and the loudspeaker. Another practical application of the transformer as an impedance matching device is its use in numerous electronic circuits as an interstage coupling device. An important circuit requirement which is satisfied by its use is that maximum power transfer occurs when the source and load resistance are equal. This is the so-called matched impedance condition.



(A) PULSE PASSING



(B) DIFFERENTIATING

Figure 3-27

2.13.4 Pulse-passing and Differentiating Transformers

In general, the characteristic of a transformer is such that it responds linearly to a sinusoidal wave. This follows from the fact that a magnetic field, varying as the current producing it, introduces a uniform variation of the induced voltage. When waveforms of other shapes are introduced into a transformer a certain element of distortion tends to develop.

In Figure 3-27(A) for example, if distortion in the output pulse is to be kept at a minimum, considerable care must be taken in choosing a primary inductance such that a relatively long time constant is formed with the resistive load R_L. When a transformer is so designed, it is usually called a pulse-passing transformer.

For special applications, transformers are sometimes used for pulse reshaping or differentiating as shown in Figure 3-27(B). In this application, the shaping of the wave is introduced by choosing values (primary inductance and R_L) that will form a short time constant. This second type of transformer is known as a "differentiating" transformer.

2.14 RELAYS

A relay in its simplest form is an electro-magnetic switch which may be used to control the operation of an electric circuit from a remote point. Its principle of operation is based on the action of a solenoid which, when energized, exerts a pulling force on any magnetizable object in the path of its magnetic field. The device

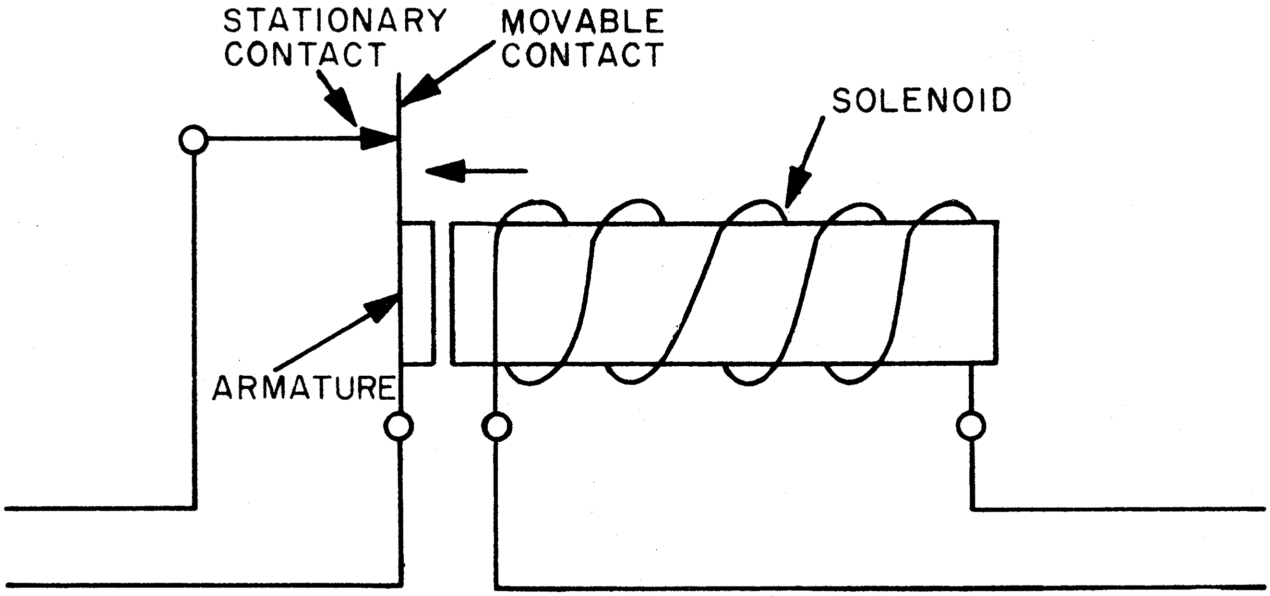


Figure 3-28

is shown schematically in Figure 3-28.

Relays are divided into two broad classes: general purpose and special purpose. Either group is essentially similar in design and construction and contains 3 basic parts: a magnetic structure, a winding or coil, and one or more contacts. Any variation in types usually involves sensitive adjustments, close marginal operation, or severe contact requirements.

The general purpose relay usually serves only as a remotely controlled switch. The construction is usually of long coil design, which permits the coil to be wound to a high resistance and to be suitable for continuous operation. In most cases the general purpose relays are given standard adjustments with no emphasis placed on operate or release time. The special purpose relays, on the other hand, are required to operate in a particular manner or under special conditions, and they include all types used to meet requirement outside the capabilities of the general purpose relay.

Relays may have one or several coils. When the relay has a single coil, the coil is used both to energize (pick up) the relay and to hold the relay energized (hold). Sometimes the relay may have a hold coil as well as the pickup coil. In this case, one coil is used to pick up the relay and the other is used to hold it. Another form of a two coil relay is a latching relay. This relay has a pickup coil which is used to energize the relay, and a mechanical latch which is used to hold the relay closed.

A latch coil is then used to unlatch the relay.

The movable contacts of a relay constitute a pole of the relay. Normally closed (abbreviated NC) is a combination of a stationary contact and a movable contact which are engaged when the coil is de-energized. When a stationary contact and a movable contact are engaged upon energizing the coil, they are called normally open contacts(abbreviated NO). A movable contact engaging a stationary contact when the coil is energized and engaging another stationary contact when the coil is de-energized is known as transfer or double-throw contact (abbreviated DT). A contact transfer in which the movable contact touches the normally open contact before leaving the normally closed contact during the transfer action is known as bridging. Therefore, the circuit of the moving contact is never completely opened. Bridging contacts are also called make-before-break contacts.

Relays that are designed to have a time-delay in operating or releasing are classified as slow operate and slow release respectively. Relays that are operated on a fairly high current value and stay operated with a considerably reduced current are known as marginal relays. Sensitive relays are used for circuit protection in instruments and low current circuits.

2.14.1 Types of Relays

2.14.1.1 Wire Contact Relays

Wire contact relays are considered as high speed relays with an operate time of 3 to 6 microseconds.

Operate time for normally open contacts is the total elapsed time from the instant the coil is energized until the contacts are closed (or open for normally closed contacts) and all contact bounce has ceased. These relays are of the multi-pole (up to 12 contacts), double-throw type of relay. They usually are used as 4 pole, 6 pole, or 12 pole relays. Although it is possible for these wire contact relays to take some arcing (very little) when used in signal circuits they are not used in high current operation. The contacts are made of spring wire.

2.14.1.2 Duo Relays

Duo relays are considered as medium speed relays with an operate time of 5 to 9 microseconds. The duo relays, like the wire contact relays, are of the multi-pole type. These relays have some current interrupting capacity and are used in the IBM card machines and in power control circuits. The contacts are made of steel strips with tungsten or silver alloy contact points. Occasionally these relays may be sealed.

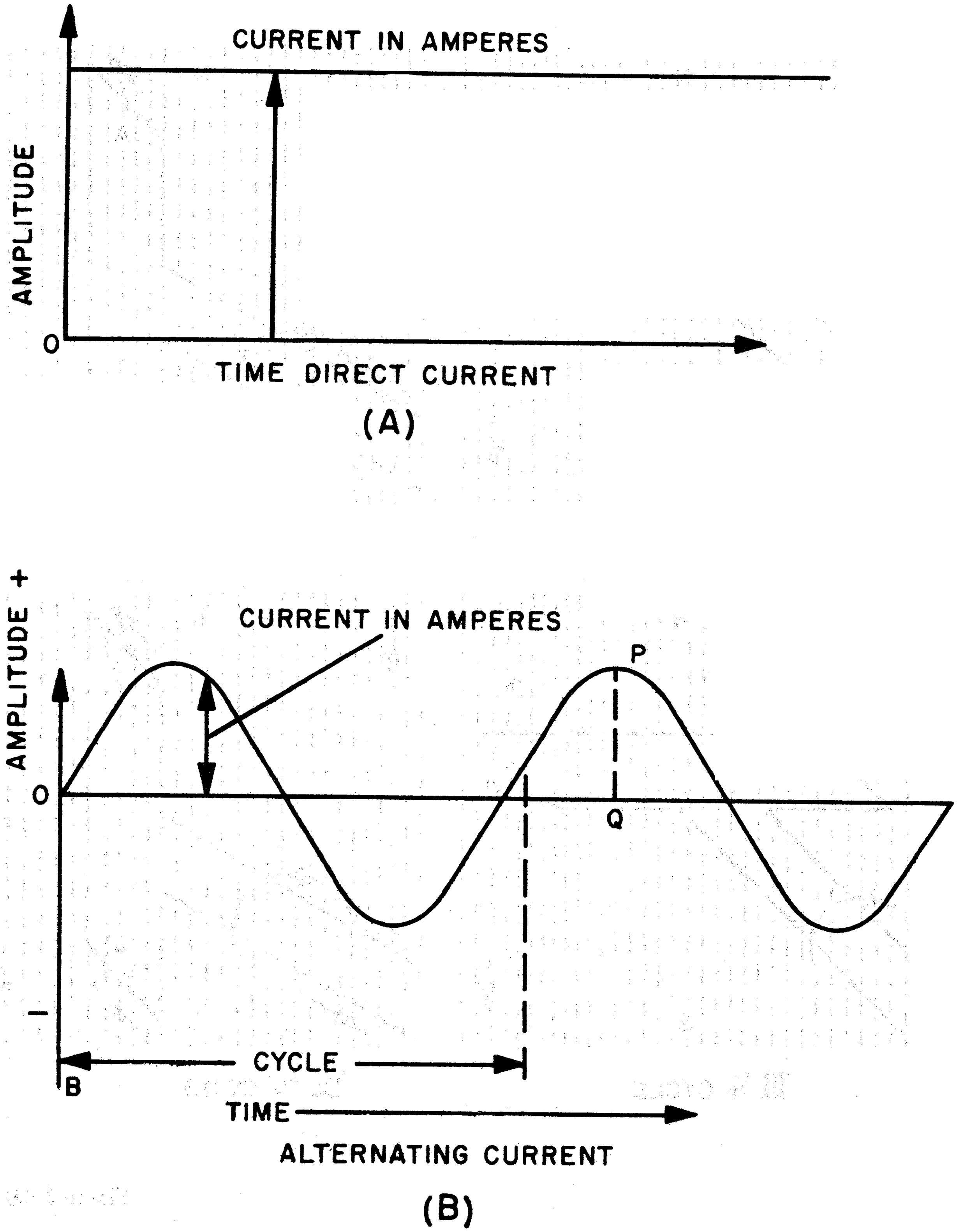

CURRENT IN AMPERES
AMPLITUDE
O
TIME DIRECT CURRENT
(A)
AMPLITUDE +
CURRENT IN AMPERES
P
O
Q
–
CYCLE
B
TIME
ALTERNATING CURRENT
(B)

Figure 3-29

PART 3

CHAPTER 3

ALTERNATING CURRENT

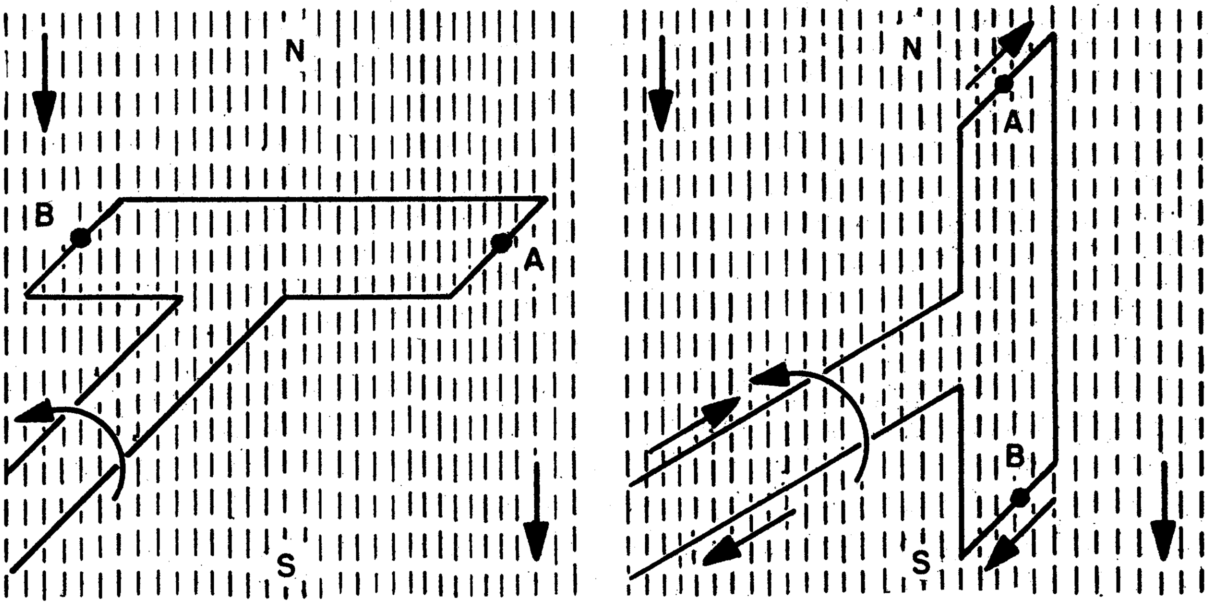
3.1 GENERAL

The source of power applied to the circuits studied in Chapter 1 is assumed to be a battery. As noted in that chapter, electron flow through a circuit is always directed away from the negative terminal of the battery and toward the positive terminal. This is unidirectional or direct current. Devices such as the alternator or a-c generator, on the other hand, produce current which changes direction at regular intervals. This is alternating current.

The form of direct current is compared to that of alternating current in Figure 3-29. Notice that the alternating current is continuously changing in magnitude and, in addition, crosses the axis of zero current flow (i.e. changes direction) at regular intervals. This continuous state of change is important because, in general, circuits possess inductive and capacitive properties which manifest themselves only when the current is changing. The effect of inductance is to impede changes in current, while the effect of capacitance is to pass only current generated by changing voltages.

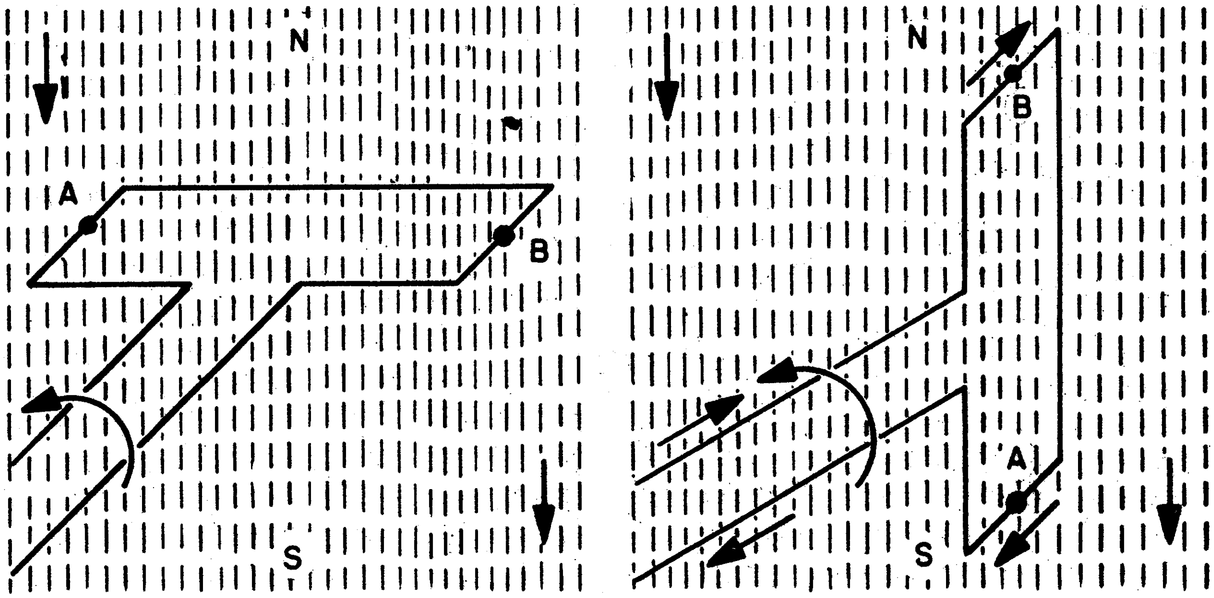
3.2 SIMPLE A-C GENERATOR

As stated in Chapter 2 of this part, a voltage can be induced in a conductor by moving that conductor through a magnetic field. This phenomenon is the basis of operation



I START

II ¼CYCLE



III ½ CYCLE

IV ¾ CYCLE

Figure 3-30

of the generator which is the chief source of electrical power in the world today. Generators are discussed in some detail in Chapter 4 of this Part. However, at this point, it is convenient to introduce a simple generator in order to facilitate the study of alternating current.

The simple generator illustrated in Figure 3-30 comprises a single loop of wire arranged so that it can be rotated through the magnetic field existing between the north and south poles of a permanent magnet. If the loop is part of a closed electrical circuit and if it is caused to rotate at a uniform rate through the field of the magnet, then an alternating current of the form shown in Figure 3-29 flows through that circuit.

To understand the form of the current in the generator loop, consider a single revolution (cycle) of the loop. Assume that the loop is moving counterclockwise and that at the start of the cycle (time $t = 0$) it is passing through the position shown in Figure 3-30-I. At this instant, the loop is moving parallel to the lines of force, so that no lines are cut by the conductor. Thus, no voltage is induced in the loop and no current flows through it. At time $t = 1/4$, the loop is passing through the position shown in Figure 3-30-II. At this instant, it is moving perpendicular to the lines of force so that it is cutting a maximum number per unit time. Thus, the induced voltage through the loop and hence the current is at a maximum. At any time between $t = 0$ and $t = 1/4$, the loop motion can be separated into two components, a decreasing component parallel to the lines of force and an increasing component perpendicular to the lines of force. Thus the number of lines cut per unit time, which is 0 at t = 0, increases continuously between t = 0 and t = 1/4, reaching a maximum at the later instant. Between t = 1/4 and t = 1/2, on the other hand, the component of loop motion parallel to the lines of force is increasing, while the component perpendicular to the lines of force is decreasing until at t = 1/2 the loop is again moving parallel to the lines of force (i.e. the perpendicular component of loop motion is 0). Thus, between t = 1/4 and t = 1/2, the number of lines cut per unit time decreases continuously until it reaches 0 at t = 1/2. Since the induced voltage, and hence the induced current, in the loop is directly proportional to the number of lines cut per unit time, the current rises from 0 at t = 0, passes through a maximum at t = 1/4 and declines again to 0 at t = 1/2. Throughout this half cycle, the direction of current through the loop remains the same, since the side of the loop labeled A is moving from right to left through the field while the side labeled B is moving from left to right through the field.

Between t = 1/2 and t = 3/4, the number of lines cut per unit time again increases reaching a second maximum at t = 3/4. Between t = 3/4 and t = 0 (the beginning of the next cycle), the number of lines cut per unit time again decreases to 0. Thus, the current rises from 0 at t = 1/2 to a maximum at t = 3/4 and falls to 0 at t = 0 of the next cycle. In magnitude, this second half cycle is identical to the first half

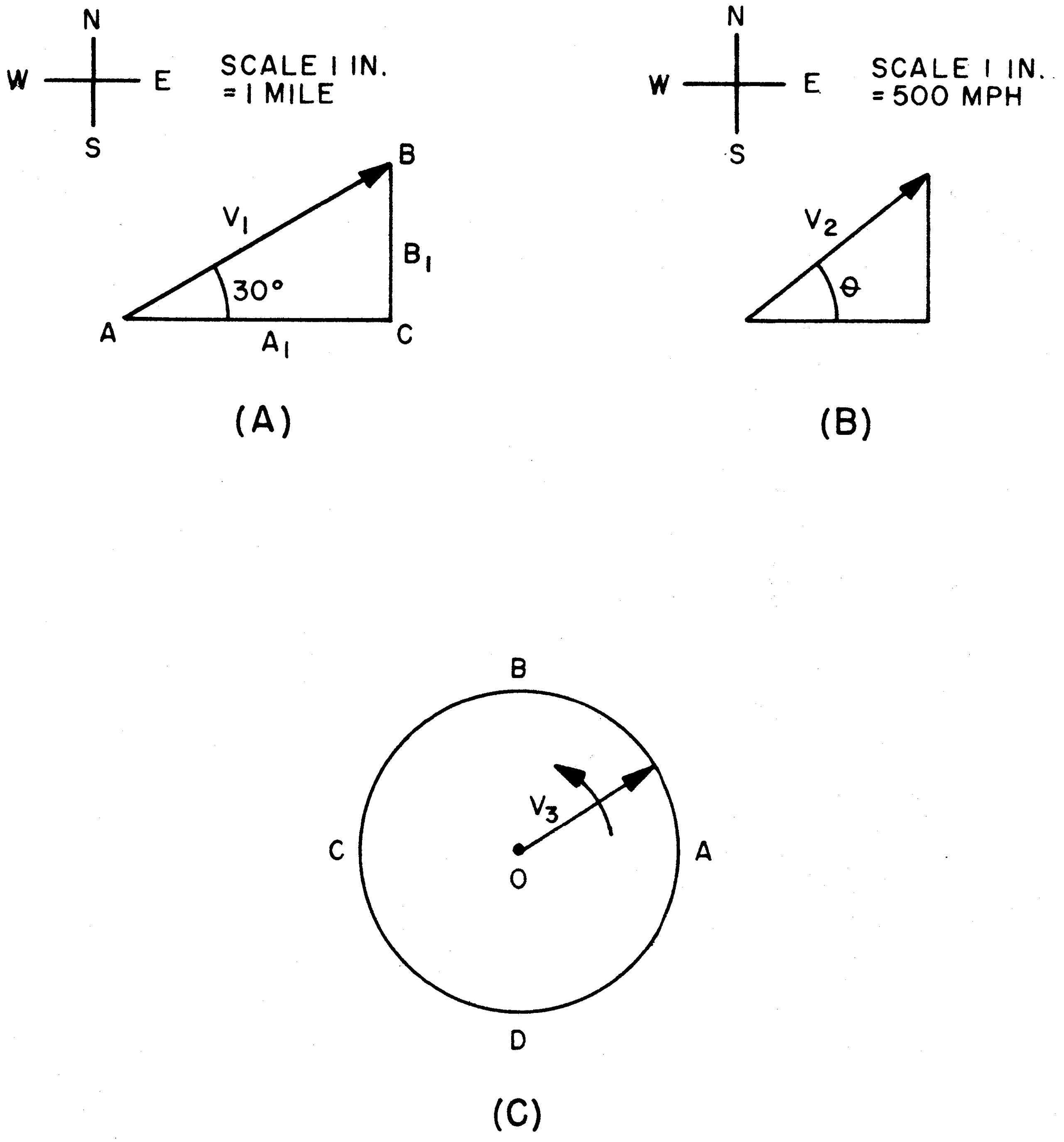

N
W
E
S
SCALE 1 IN.
= 1 MILE
B
V_1
B_1
30°
A
A_1
C
(A)
N
W
E
S
SCALE 1 IN.
= 500 MPH
V_2
θ
(B)
B
V_3
C
O
A
D
(C)

Figure 3-31

cycle. However, the direction of current during the second half cycle is opposite to that during the first half cycle. This follows from the fact that during the second half cycle, the A side of the loop is cutting across the magnetic field from left to right and the B side of the loop is cutting across the field from right to left which is exactly opposite to the situation existing during the first half cycle.

3.3 VECTORS

A vector is a straight line which is used to represent direction as well as magnitude. Consider, for example, vector V_1 of Figure 3-31-(a). This is a line two inches long which makes an angle of 30° with the horizontal. The arrowhead indicates that the line is pointing upward and to the right rather than downward and to the left. The coordinate system indicates that the vector is to be orientated with respect to north, south, east and west. The scale (1 inch = 1 mile) indicates that V_1 represents a distance. Specifically, V_1 represents a distance of two miles in a direction 30° to the north of east.

Vector V_1 is shown as the hypotenuse of right triangle ABC having side a_1 parallel to the horizontal (x) axis and side b_1 parallel to the vertical (y) axis. Side a_1 is said to be the x-axis component, or the projection on the x axis, of V_1, while side b_1 is said to be the y-axis component, or the projection on the y axis, of V_1. In terms of trigonometry:

$$b_1 = V_1 \sin 30^\circ$$

$$a_1 = V_1 \cos 30^\circ$$

It is not necessary to show the components of a vector. For example, vector V_2 of Figure 3-31-(b) is shown without components. This vector is to be interpreted with respect to a north-south-east-west coordinate system just as is vector V_1 of Figure 3-31-(a). However, the scale (1 inch = 500 mph) indicates that V_2 (which is 1-1/2 inches long) represents a speed of 750 mph in a direction θ° to the north of east.

Figure 3-31-(c) illustrates a vector which is assumed to be rotating in a counterclockwise direction. The length of the vector itself remains constant as it moves; however, the lengths of its x-axis and y-axis components change as it moves. This corresponds to the fact that these components are functions of the angle θ between the x-axis and the vector. Specifically:

$$a = V \cos \theta$$

$$b = V \sin \theta$$

where a is the x-axis component and b is the y-axis component of vector V.

Referring to Figure 3-31-(c), when vector V_3 is pointing at A, $b = 0$ (since $\sin 0^\circ = 0$). When V_3 is pointing at B, $b = V_3$ (since $\sin 90^\circ = 1$). When V_3 is pointing at C, $b = 0$ (since $\sin 180^\circ = 0$). When V_3 is pointing at D, $b = - V_3$ (since $\sin 270^\circ = -1$). Consider the behavior of b through one cycle of rotation of V_3: it is 0 at A which is taken as the starting point of the cycle, rises to a positive maximum at B, falls to 0 again at C, rises to a negative maximum at D and finally returns to 0 as V_3 reaches A and begins a new

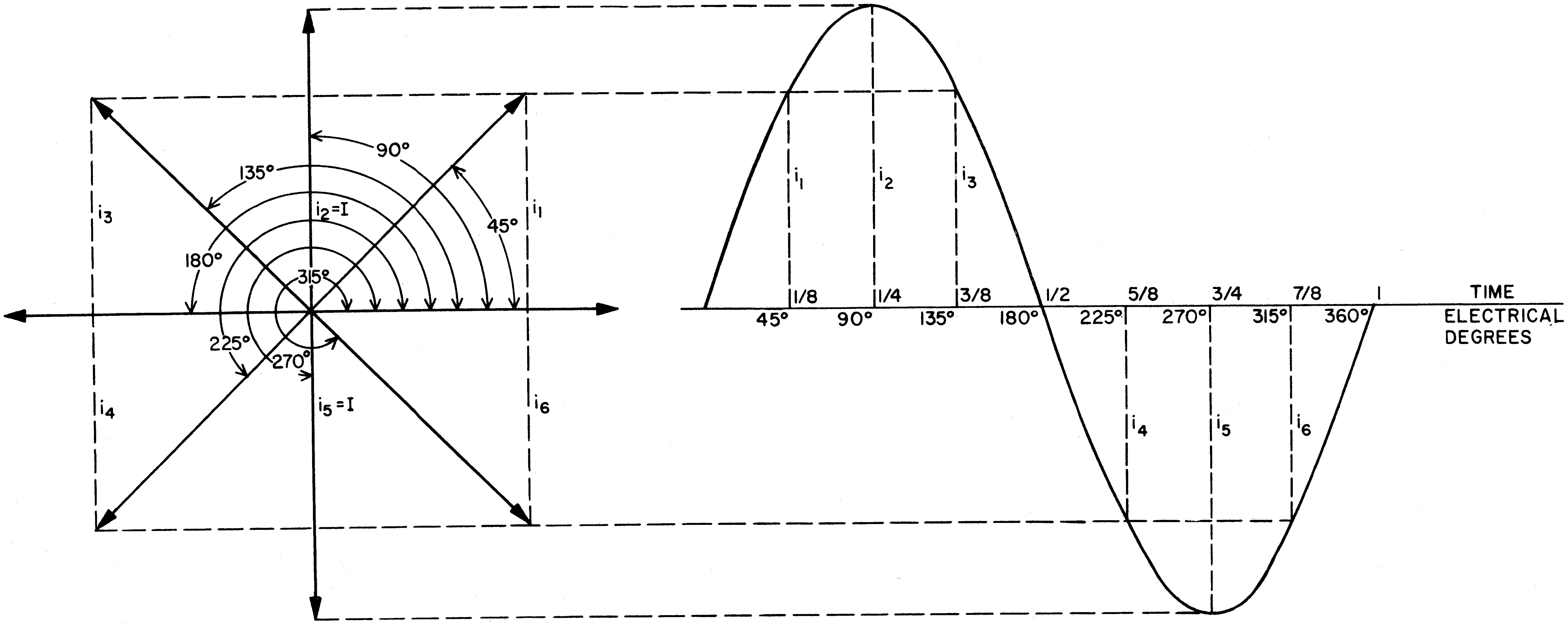

90°
135°
45°
180°
315°
225°
270°
i2=I
i5=I
i1
i3
i4
i6
i1
i2
i3
i4
i5
i6
1/8
1/4
3/8
1/2
5/8
3/4
7/8
1
TIME
45°
90°
135°
180°
225°
270°
315°
360°
ELECTRICAL DEGREES

Figure 3-32

cycle. A comparison of this behavior of b through a cycle of rotation of V_3 with the behavior of the induced current through a cycle of rotation of the simple generator discussed in the preceding section, reveals that the two are analogous. In fact, if $V_3 = I$, where I is the maximum value of induced current, then the y-axis component of V_3 represents the instantaneous value of the induced current. Thus, in terms of trigonometry:

$$i = I \sin \theta$$

where:

i = instantaneous value of current

I = maximum value of current

θ = point of cycle being considered

For example, if I = 3 amperes and $\theta = 30^\circ$, then:

$$\begin{aligned} i &= I \sin \theta \\ &= (3)\ (\sin 30^\circ) \\ &= (3)\ (1/2) \\ &= 3/2 \text{ amp} \end{aligned}$$

3.4 SINE WAVE

In Figure 3-29, the waveform of an alternating current such as is produced by an a-c generator is illustrated. In Section 3.3, it is explained that this current can be represented by a rotating vector. It remains to show the relationship between the waveform of Figure 3-29 and the rotating vector of Figure 3-31-(c). This relationship is illustrated in Figure 3-32. Reference to this figure reveals that the waveform is actually a plot of the instantaneous value of the y-axis component of the vector (i.e. the maximum value of current, I) against electrical degrees or against time. Since the angular velocity of the vector is assumed to be constant, there is a fixed relationship between time and electrical degrees so that both can be shown along the same axis. Since the y-axis component of the vector which the waveform represents is $I \sin \theta$, the waveform is properly called a sine wave.

The concepts of frequency, amplitude, effective value and phase difference can now be developed in terms of sinusoidal alternating current.

3.4.1 Frequency

Frequency is the measure of speed of alternation of current. It is stated in cycles per second. Thus, if ten complete cycles of alternating current are completed in three seconds, then the frequency of the current is 10/3 cycles per second (cps).

3.4.2 Amplitude

Since the magnitude of alternating current is continuously changing, it is convenient to use the maximum value, I, as a measure of magnitude. This measure is called the amplitude of the current.

As long as the current is sinusoidal in form, the instantaneous value of the current at any time in the cycle can be found by solving the equation:

$$i = I \sin \theta$$

which is developed in Section 3.3 above.

3.4.3 Effective Value

Assume that a direct current of I amperes produces a certain amount of heat in a given resistance per unit time. It is natural to inquire whether an alternating current of amplitude I amperes will produce the same result. It turns out that it will not. It is, therefore, convenient to have another measure of the value of alternating current in terms of which I amperes of alternating current can be expected to produce the same heating effect as I amperes of direct current. This measure is called the effective or rms value of alternating current. An effective or rms alternating volt is that alternating potential which drives an effective current of 1 ampere through a pure resistance of 1 ohm. Line voltage is usually specified in terms of its rms or effective value rather than in terms of it amplitude.

3.4.4 Phase

Two currents can be of the same frequency and yet change direction at different times. For example, if 0° for current i_1 corresponds to 90° for current i_2, then i_1 is passing through 0 just as i_2 is passing through a maximum. In this case, the two currents are said to be 90° out of phase. Moreover, i_1 is said to lag i_2 by 90°, while i_2 is said to lead i_1 by 90°.

An alternating current results when an alternating voltage is impressed upon a closed circuit. The alternating voltage, like the current which results from it, is sinusoidal in form and can be represented by a rotating vector or a sine wave.

Moreover, if the maximum impressed voltage is E, then the instantaneous voltage, e, is given by the equation:

$$e = E \sin \theta.$$

If an alternating voltage is impressed upon a purely resistive circuit, then the resultant current is in phase with the impressed voltage; that is, current and voltage pass through 0 simultaneously, pass through a positive maximum simultaneously, pass through 0 a second time simultaneously and pass through a negative maximum simultaneously. In terms of vectors, the direction of the current vector coincides with that of the voltage vector.

When an alternating voltage is impressed upon a circuit containing inductance or capacitance as well as resistance then the voltage and current are no longer in phase. In this case, the angle between the voltage vector and the current vector is called the phase angle.

3.5 INDUCTANCE

As indicated in Section 2.13 of this part, a change in the magnitude of current flowing through a conductor causes an expansion or contraction of the magnetic field around the conductor. Some of the expanding or contracting lines of flux cut across the conductor itself, generating a counter-emf of self-induction which always opposes the change in current that causes it. The magnitude of the counter-emf that is induced in any circuit by a unit rate of change of current is called the inductance of the circuit. Inductance is represented by the symbol L and is expressed in henries. A circuit

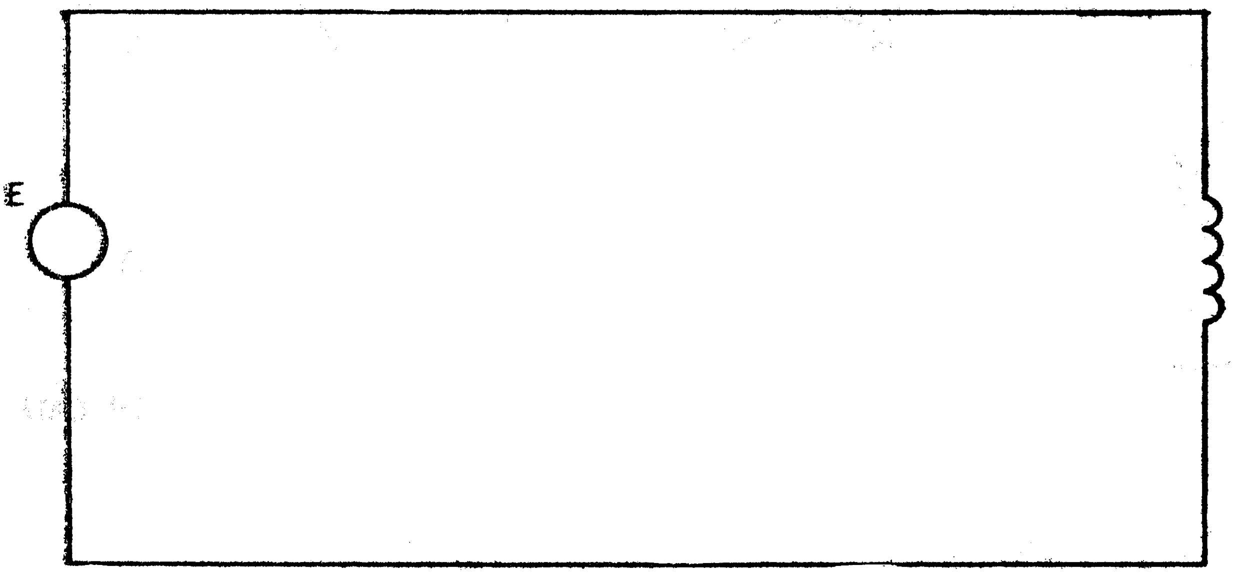


Figure 3-33

has an inductance of 1 henry when a uniform rate of change of current of 1 ampere per second produces a counter-emf of 1 volt.

The magnitude of the counter-emf that is induced by a unit rate of change of current, in a circuit, that is the inductance of the circuit, depends upon the number of lines of flux that cut across the conductor per unit time per unit rate of change of current. This, in turn, depends upon the physical arrangement of the conductor and the character of the associated magnetic circuit. When the conductor is coiled to form a solenoid, then the expanding or contracting lines of flux associated with each turn cut across adjacent turns inducing a counter-emf in them. Since a unit rate of change of current induces a voltage in each turn of a solenoid and since all the turns are in series, it follows that the total counter-emf induced in a solenoid is proportional to the number of turns. Moreover, if the conductor is wound around an iron core then the reluctance of the magnetic circuit is greatly reduced so that the flux density per unit rate of change of current is increased with a consequent increase in the inductance of the circuit.

3.6 INDUCTIVE REACTANCE

All practical circuits contain some ohmic resistance. However, for the purpose of evaluating the effects of inductance, it is convenient to consider the theoretical case of a pure inductance connected in series with a source of alternating emf. Such a circuit is shown schematically in Figure 3-33.

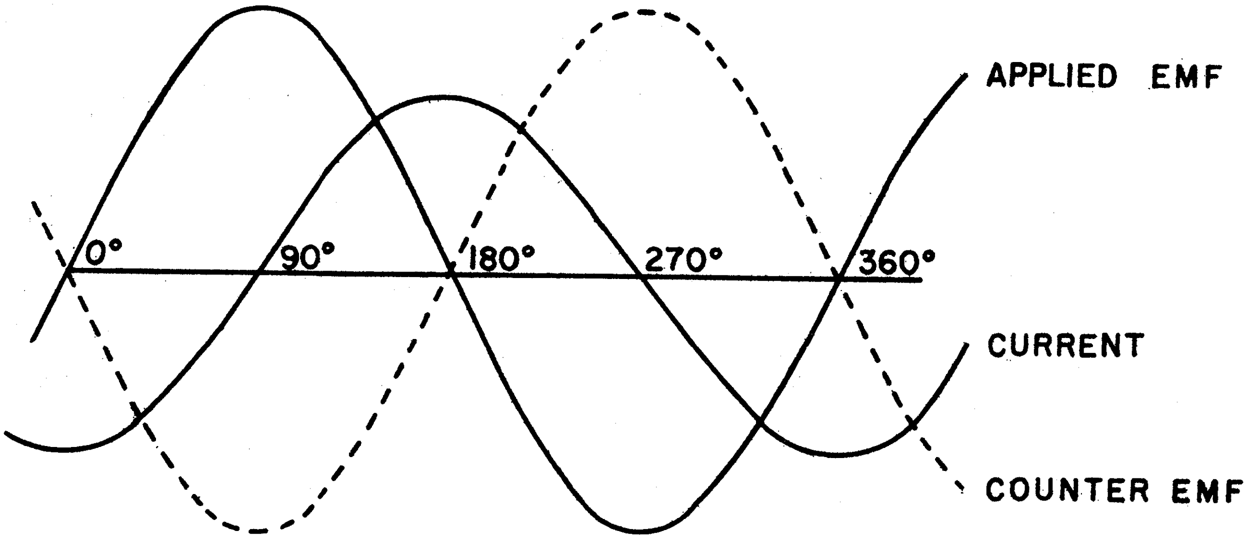


Figure 3-34

Since there is assumed to be no resistive opposition to the flow of current in the circuit of Figure 3-33, an infinite current should flow in one direction during one half cycle and in the other direction during the other half cycle. However, any change in the magnitude of current induces a counter-emf which opposes that change. The rate of change of current at any instant is just sufficient to induce a counter-emf which is equal and opposite to the instantaneous value of the applied emf. From this, it follows that the rate of change of current at any instant is directly proportional to the instantaneous value of the applied voltage and inversely proportional to the inductance of the circuit.

The phase relationship between applied emf, counter-emf of self-induction and current for the circuit of Figure 3-33 is shown in Figure 3-34. Referring to the latter figure, notice that the current reaches a maximum value just as the applied voltage passes through zero. This follows from the fact that the current tends toward an infinite value in one direction as long as the applied emf is positive, and tends toward an infinite value in the opposite direction as long as the emf is negative. Thus, each time that the applied emf passes through zero, the increase of the current in one direction ceases and a decrease begins (since the current is now tending toward an infinite value in the opposite direction). As the applied voltage reaches a maximum, the current passes through zero. Thus, the current is 90° out of phase with the applied emf. In terms of current and voltage vectors, the current lags the applied emf by 90°.

The amplitude of the current is directly proportional to the rate of change of current; that is, the faster the current increases, the greater is the maximum value it attains. It has already been noted that the rate of change of current, in turn, is inversely proportional to the inductance of the circuit. Thus the amplitude of the current is inversely proportional to the inductance. At the same time, the amplitude of the current is inversely proportional to the frequency of the applied emf; that is, the higher the frequency of the applied emf, the shorter is each interval during which the current increases. These relations can be restated in terms of opposition to current offered by an inductive load, as follows:

$$X_L = 2\pi f L \qquad (15)$$

where:

X_L = inductive reactance in ohms

f = frequency of the applied emf in cycles per second

L = inductance in henries

Here, 2π is a conversion factor arising from the fact that equation (15) expresses inductive reactance in terms of the angular velocity of the voltage vector in radians per second, rather than directly in terms of frequency, (2π radians = 360°).

Referring again to Figure 3-34, notice that during the portions of the applied voltage cycle when the voltage is increasing (as, for example, between 0° and 90°), the current waveform is on the opposite side of the zero axis from the applied voltage waveform. This corresponds to the fact that current is flowing

through the load way from positive terminal of the generator and toward the negative terminal (i.e. power is being returned from the inductance of the load to the source). On the other hand, during portions of the applied voltage cycle when the voltage is decreasing, (as, for example, between 90° and 180°), the current waveform is on the same side of the zero axis as the applied voltage waveform. This corresponds to the fact that current is flowing through the load away from the negative terminal of the generator and toward the positive terminal (i.e. power is being taken from the source by the inductive load). Thus, power is transferred from the source to the load for half of each cycle and from the load to the source for the other half of that cycle. The result is that the total power dissipated by the purely inductive load is 0 watts. Since any practical circuit contains some ohmic resistance, this situation is theoretical.

3.6.1 Inductances in Series and Parallel

In order to compute the reactance of inductive networks, it is necessary to know how to reduce series or parallel inductances to effective inductances. The rules in either case are similar to those governing the behavior of resistances in series and in parallel. Thus, if L_1, L_2 and L_3 are connected in series, the total inductance, L_t, is:

$$L_t = L_1 + L_2 + L_3 \tag{16}$$

Similarly, for L_1, L_2 and L_3 in parallel:

$$L_t = \frac{1}{\frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3}} \tag{17}$$

3.7 CAPACITANCE

In approaching the subject of capacitance, it is important to understand the distinction between charge and emf.

The measure of charge is the coulomb. Specifically, 1 coulomb = 6.3×10^{18} electrons. A positive charge of 1 coulomb is a deficiency of that number of electrons, while a negative charge of 1 coulomb is an excess of that number of electrons.

The measure of emf is the volt. A constant emf of 1 volt drives a current of 1 ampere through a resistance of 1 ohm. EMF depends upon concentration of charge. A charge of 1 coulomb distributed evenly over the surface of the earth would create no significant potential difference between the earth and the surrounding atmosphere. On the other hand, in a practical circuit, a charge on the order of .005 coulomb can easily be concentrated so that it exerts an emf of 1200 volts.

The ability of a circuit to accommodate charge is called its capacitance. A given quantity of charge is more concentrated when stored in a circuit of small capacitance than it is when stored in a circuit of large capacitance. Thus, a given charge exerts a larger emf when stored in a circuit of small capacitance than it does when stored in a circuit of large capacitance. This relationship is expressed by the following equation:

$$C = Q/E \tag{18}$$

where

C = capacitance in farads

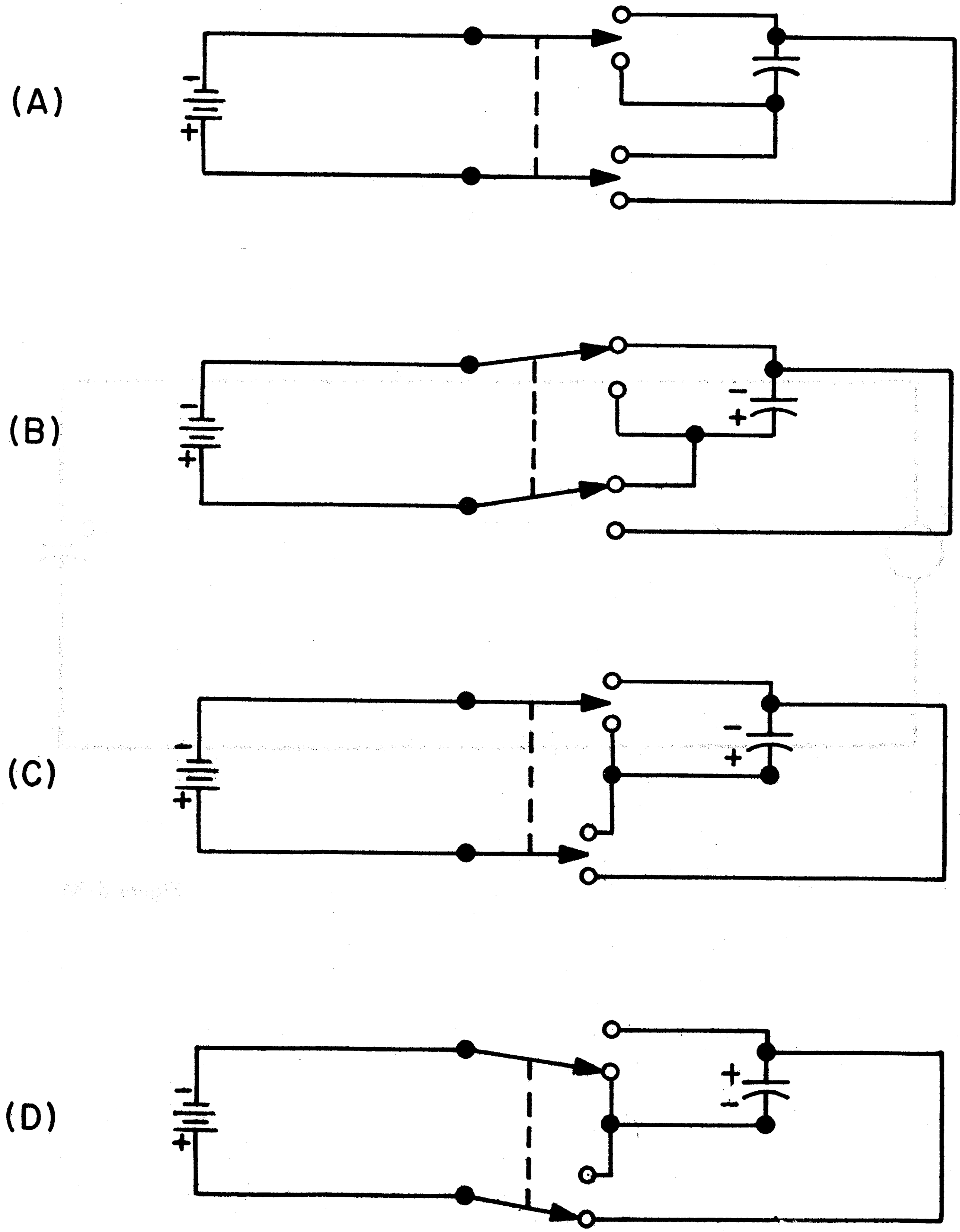


Figure 3-35

Q = charge in coulombs

E = emf in volts

A conductor is characterized by its ability to yield or accept electrons. Thus, every circuit has a certain amount of distributed capacitance. At the same time, there are many reasons for concentrating capacitance at various points in a circuit. Elements which provide this concentration of capacitance are called capacitors.

A simple form of capacitor comprises two parallel plates separated by an insulator such as air, mica or glass. The insulating material between the plates is called the dielectric.

The capacitance of any particular capacitor depends upon the electrical characteristics of the materials out of which the plates and dielectric are formed and also upon the dimensions of plates and dielectric. For a capacitor constructed of any particular materials, capacitance is directly proportional to the area of the plates and inversely proportional to the distance between them.

An arrangement including a capacitor, a switch and a battery is shown in Figure 3-35. Part (a) of the figure shows the switch open so that the capacitor is isolated from the battery. Here, it is assumed that the capacitor has not been previously connected to the battery, so that there is no potential difference between its upper and lower plates. In part (b), the switch has been thrown to the position which connects the negative terminal of the battery to the upper plate of the capacitor and the positive terminal of the battery to the lower plate of the capacitor. Under this condition, current flows through the circuit until the potential difference between the plates of the capacitor is equal to the emf of the battery. No current actually flows through the capacitor; electrons simply move onto the upper plate and move away from the lower plate. When the current has ceased, the capacitor is said to be fully charged. The quantity of charge stored by the fully charged capacitor is, by equation (18);

$$Q = CE$$

If the switch is now opened as shown in part (c) of the figure, there is no complete current path through which the charge can distribute itself. Thus, the potential difference between the capacitor plates remains constant. In part (d) of the figure, the switch has been positioned so as to connect the negative terminal of the battery to the lower plate of the capacitor and the positive terminal of the battery to the upper plate of the capacitor. Under this condition, the capacitor first discharges through the battery and then charges in the opposite direction as indicated.

If the battery and switch of Figure 3-35 are replaced by an a-c generator, then alternating current flows through the circuit as the potential difference between the capacitor plates continuously changes in response to the changing applied input voltage. By first charging in one direction, then discharging, then charging in the other direction, the capacitor acts as a conductor even though no current actually flows through the dielectric between its plates.

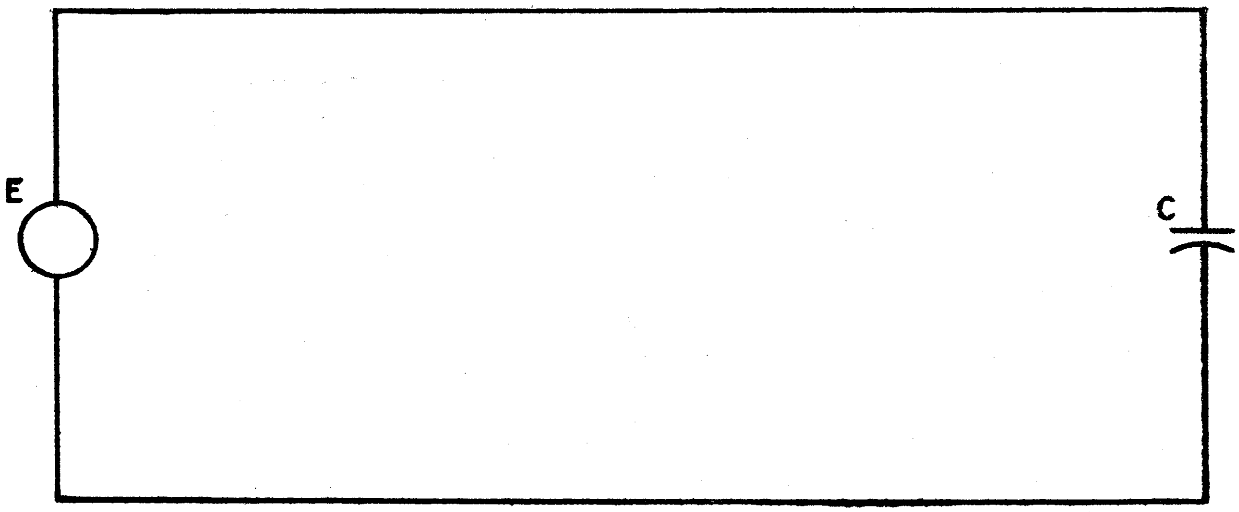


Figure 3-36

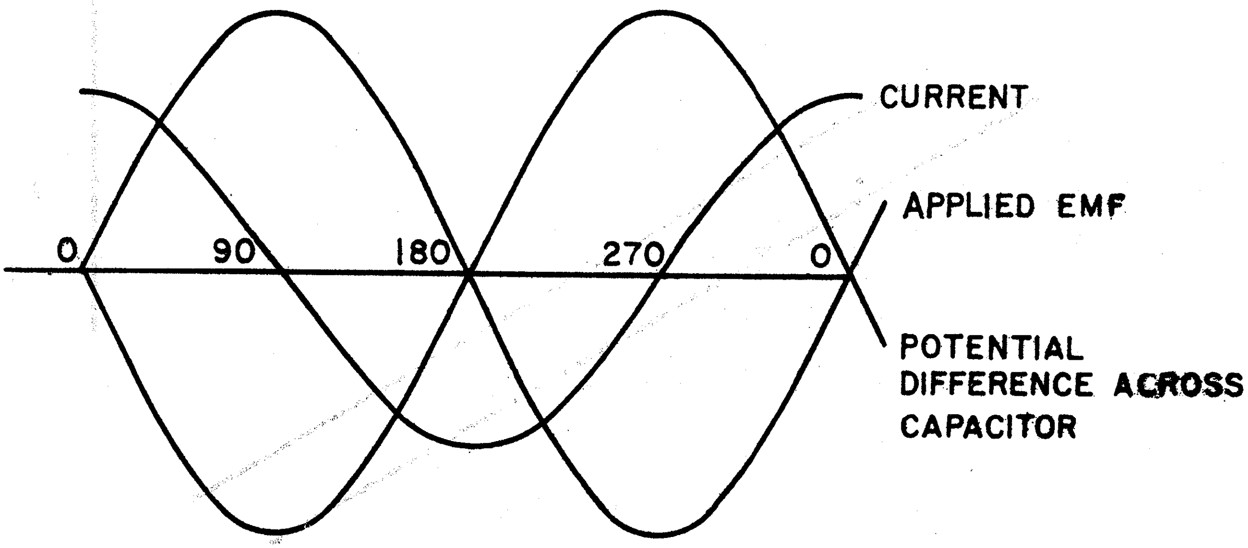

CURRENT
APPLIED EMF
0
90
180
270
0
POTENTIAL DIFFERENCE ACROSS CAPACITOR

Figure 3-37

3.8 CAPACITIVE REACTANCE

In order to evaluate the behavior of current in a capacitive circuit, it is convenient to consider the theoretical case of an alternating source of emf in series with a purely capacitive load. Such a circuit is shown schematically in Figure 3-36. The phase relations between applied emf, potential difference across the capacitor and current, for the circuit of Figure 3-36, are shown in Figure 3-37.

Since there is assumed to be no resistive opposition to the flow of current in this circuit, any emf should cause an infinite current flow. However, flow of current through the circuit creates a potential difference across the capacitor which opposes the emf of the source. Thus, the value of current which flows in the circuit is just sufficient to maintain a potential difference across the capacitor which is equal and opposite to the applied emf. From this it follows that the magnitude of the current at any instant is proportional to the rate of change of applied emf at that instant. Since the rate of change of voltage is zero when the voltage is maximum and maximum when the voltage is zero, the current is zero when the applied voltage is maximum and maximum when the applied voltage is zero (i.e. the current is 90° out of phase with the applied voltage). Moreover, when the applied voltage is increasing (as, for example, between 0° and 90°), then the capacitor is charging (i.e. it is taking power from the source) so that current is flowing through the load from the instantaneously negative terminal of the generator to the instantaneously positive terminal. Thus, in Figure 3-37, the applied voltage and current waveforms appear on the same side of the zero axis during the parts of the voltage cycle when the applied voltage is increasing. On the other hand, when the applied voltage is decreasing (as, for example, between 90° and 180°), then the capacitor is discharging (i.e. it is returning power to the source) so that the current is flowing through the load from the instantaneously positive terminal of the generator to the instantaneously negative terminal. In terms of current and applied voltage vectors, the current leads the voltage by 90°.

As has already been stated, the magnitude of current produced in the purely capacitive circuit is just sufficient to maintain an equality between the applied emf and the opposing potential difference across the capacitor. But, by equation (18), the potential difference across the capacitor is:

$$E = Q/C$$

Thus, the quantity of charge that must be moved (i.e. the current), per unit rate of change of potential difference, is directly proportional to the capacitance of the circuit. It has already been shown that the magnitude of current is directly proportional to the rate of change of the applied emf. These relations can be restated in terms of opposition to current offered by a capacitive load, as follows:

$$X_C = \frac{1}{2\pi f C} \tag{19}$$

where:

X_C = capacitive reactance in ohms

f = frequency of applied emf in cycles per second

C = capacitance in farads

For practical circuits, the farad is an extremely large unit. For this reason, capacitance is usually expressed in microfarads (mf) or micromicrofarads (mmf), where:

1 mf = 10^{-6} farad

1 mmf = 10^{-12} farad

As has been noted, the purely capacitive load takes power from the source when it is charging and returns power to the source when it is discharging. Since it is charging for half of each cycle and discharging for the other half and since the charging and discharging occur in response to symmetrical voltage conditions, the amount of power taken from the source is equal to the amount of power returned to the source. Thus, the total power dissipated by the purely capacitive circuit is 0 watts. Since any practical circuit contains some ohmic resistance, this situation is theoretical.

3.8.1 Capacitors in Series and Parallel

In order to compute the capacitive reactance of networks of capacitors, it is necessary to be able to compute the effective capacitance of capacitors in series or in parallel. A number of capacitors in parallel function like a single capacitor of greater plate area. Thus:

$$C_t = C_1 + C_2 + \dots + C_n \qquad (20)$$

where C_t is the total capacitance of n capacitors connected in parallel.

Connecting capacitors in series is equivalent to increasing the distance between the plates; thus, the total capacitance of n capacitors in series is less than the capacitance of any of the individual capacitors. Specifically, the equation for the total capacitance of n series-connected capacitors is:

$$C_t = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} + \dots + \frac{1}{C_n}} \qquad (21)$$

3.9 IMPEDANCE

The three parameters which define the response of a circuit to any particular applied emf are ohmic resistance, inductive reactance and capacitive reactance. The total effect of all three parameters is called impedance. In the following sections, impedance is discussed in terms of various combinations of resistance and reactance and also in terms of emf and current.

3.9.1 Impedance of Series R-C Circuits

When an alternating voltage is applied to a circuit consisting of capacitance and resistance in series the total opposition which the circuit offers to the flow of current depends upon the values of both capacitive reactance and ohmic resistance. The capacitive reactance tends to force the current to lead the voltage by an amount which is directly proportional to the ratio of capacitive reactance to ohmic resistance.

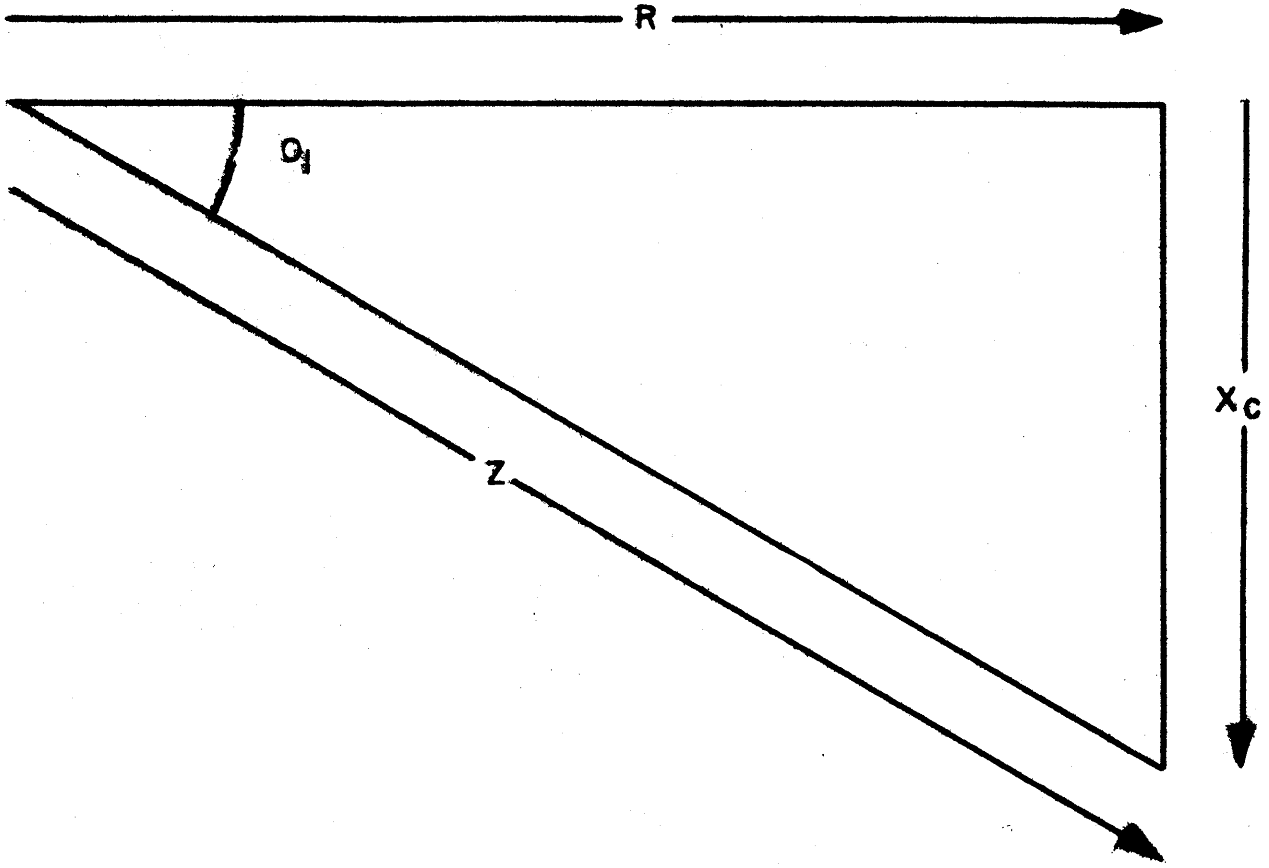


Figure 3-38

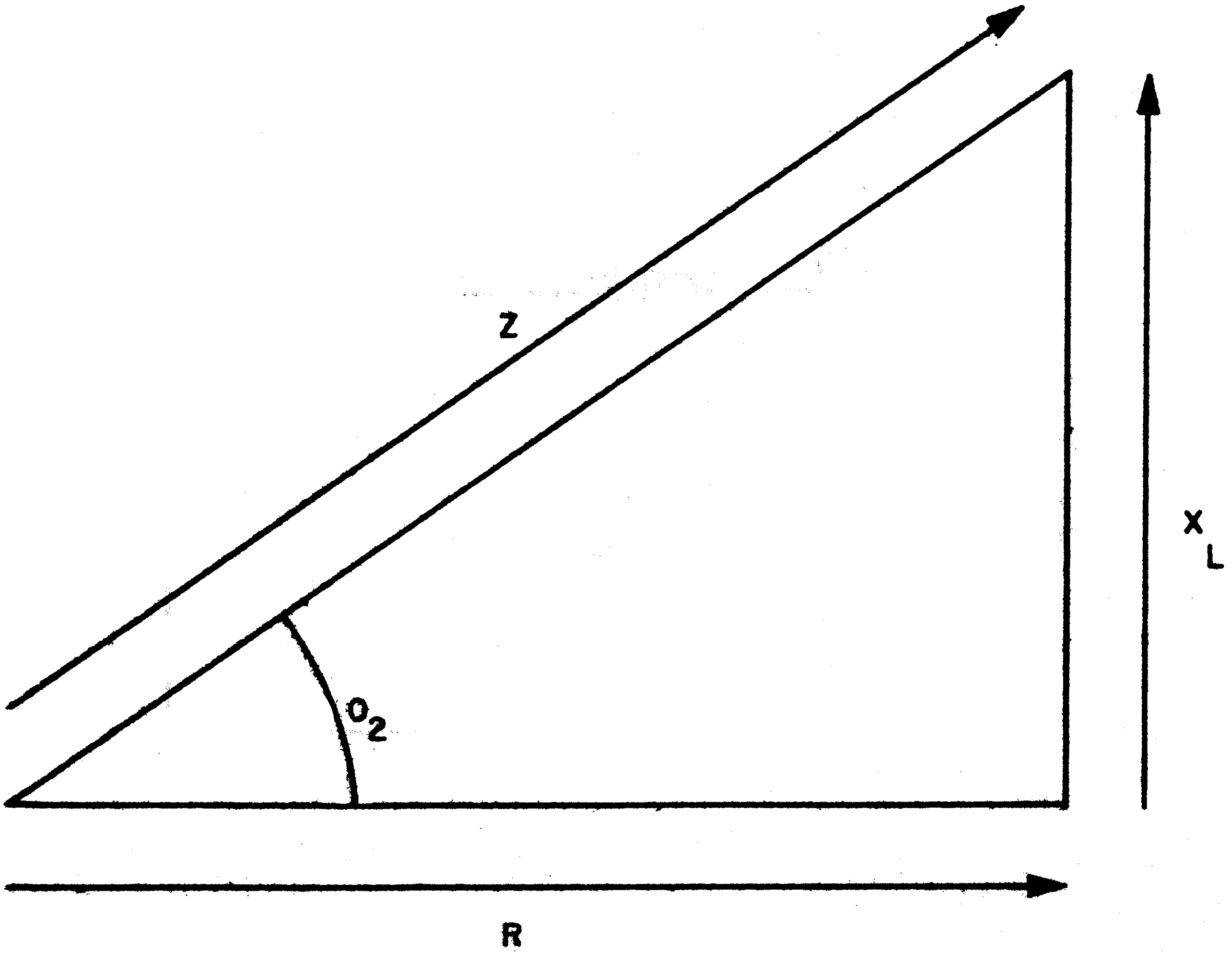


Figure 3-39

The so-called impedance triangle, shown in Figure 3-38 is useful in establishing not only the total value of the opposition or impedance offered to the flow of current, but also in establishing the angle by which the current is caused to lead the voltage. Referring to the figure, notice that the triangle has sides of length R and X_c which are at right angles to each other. By the pythagorian theorem, the third side, which represents the impedance, is given by the equation.

$$Z = \sqrt{R^2 + X_c^2} \tag{22A}$$

The angle θ_1 (called the impedance angle) defines the amount by which the current leads the voltage.

3.9.2 Impedance of Series R-L Circuits

When an alternating voltage is applied to a circuit consisting of inductance and resistance, the total opposition which the circuit offers to the flow of current depends upon the values of both inductive reactance and ohmic resistance.

An impedance triangle similar to that used for the R-C circuit is useful in establishing both the magnitude of the impedance and the angle by which the current lags the voltage. The R-L impedance triangle is shown in Figure 3-39. Comparing Figures 3-38 and 3-39, notice that X_L is shown above the resistance axis while X_C is shown below the resistance axis. This corresponds to the fact that they have opposite effects on the phase relationship between voltage and current. θ_2 defines the amount by which the current lags the voltage. The impedance of an R-L series circuit is given by the formula:

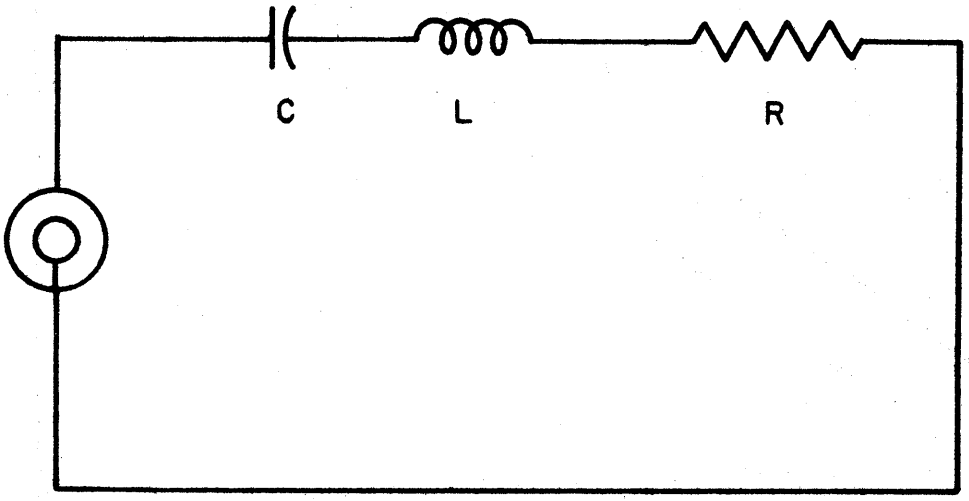


Figure 3-40

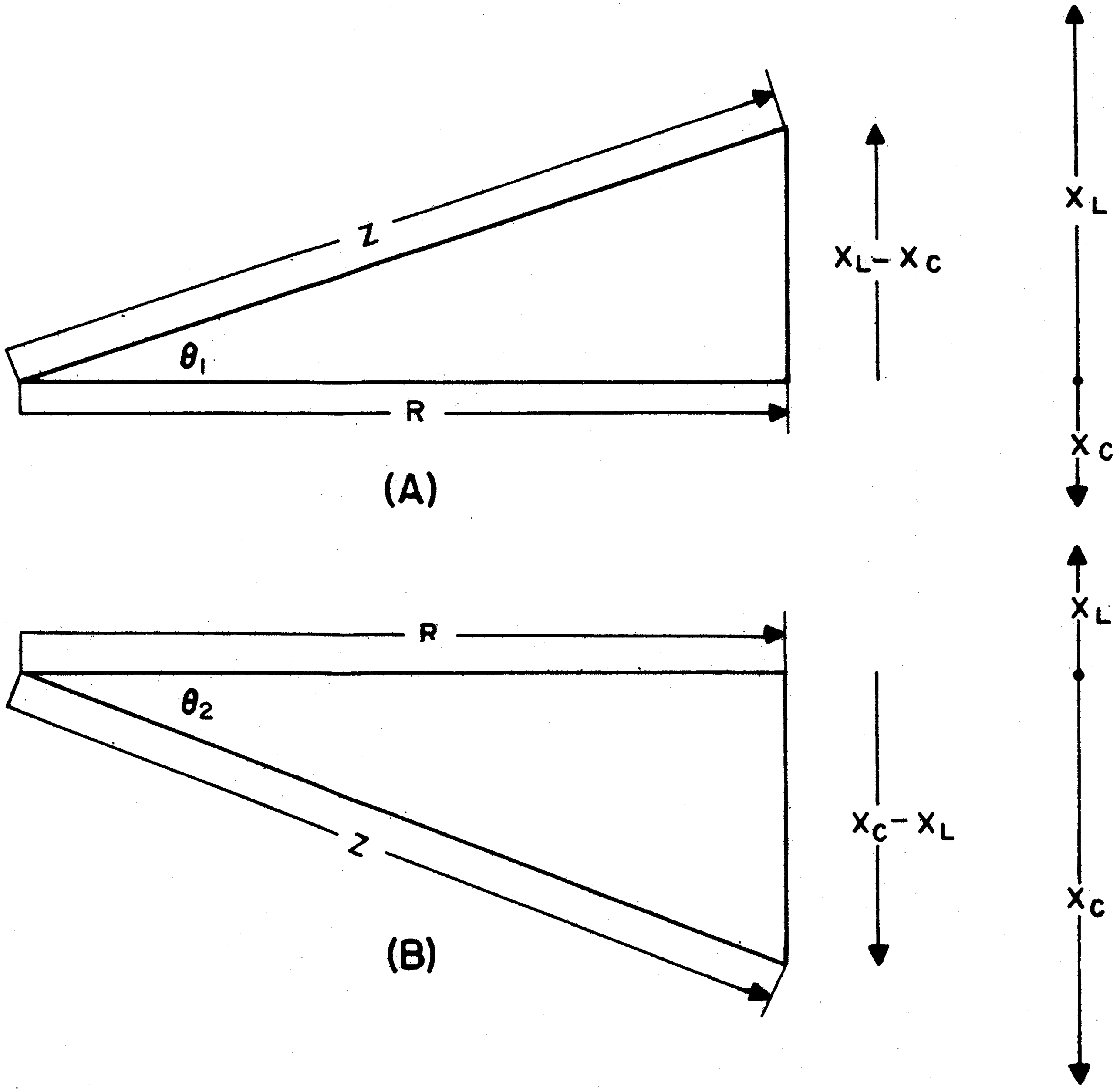


Figure 3-41

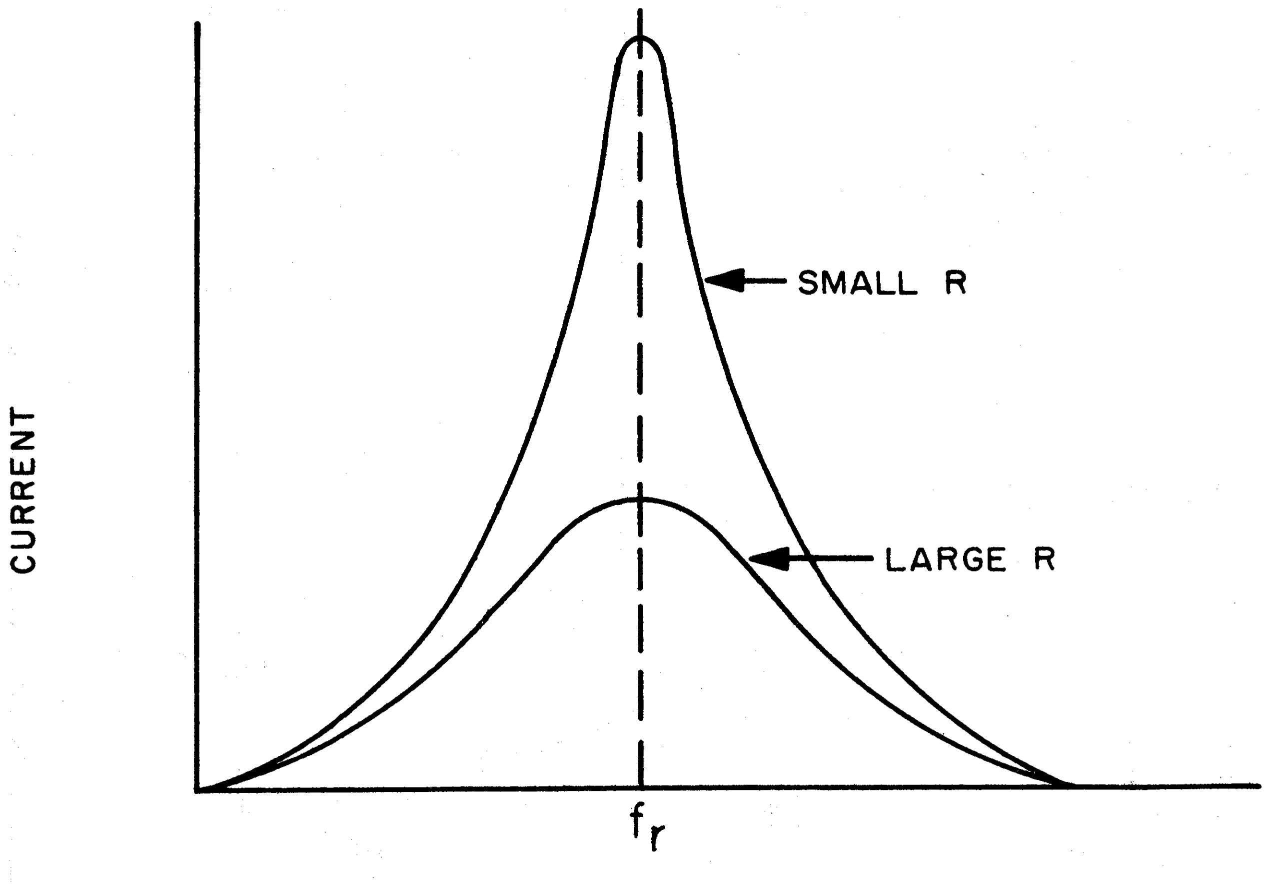


Figure 3-42

$$Z = \sqrt{R^2 + X_L{}^2} \qquad (22\ B)$$

3.9.3 Impedance of Series R-L-C Circuits

When inductance, capacity and resistance are connected in series in an alternating current circuit, as shown in Figure 3-40, the total impedance is given by the formula:

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \qquad (23)$$

An impedance triangle representing the magnitude of impedance and the angle of the lead or lag can be drawn after the capacitive reactance has been subtracted from the inductive reactance. Two such triangles are shown in Figure 3-41. Triangle (A) represents a circuit where the inductive reactance is larger than the capacitive reactance. Hence, the angle θ_1 is a lag angle and is shown above the resistance axis. Triangle (B) represents a circuit where the capacitive reactance is larger than the inductive reactance. Hence, θ_2 is a lead angle and is shown below the resistance axis.

Maximum current in a series R-L-C circuit is obtained when $X_L = X_C$. At this point, $Z = R$, and $I = \frac{E}{R}$. The circuit is said to be in series resonance for such a condition. At frequencies below resonance, the current leads the voltage because the capacitive reactance is greater than the inductive reactance while, at frequencies above the resonant frequency, the current lags because the inductive reactance is greater than the capacitive reactance. Figure 3-42 shows two resonant curves for different values of R. Since, at resonance, $X_L = X_C$,

$$2\pi fL = \frac{1}{2\pi fC}$$

Solving for the resonant frequency,

$$2\pi f = \frac{1}{2\pi \quad LC} \tag{24}$$

3.9.4 Impedance of Parallel Combinations of L,C and R

Consider a theoretical parallel circuit consisting of a purely capacitive branch and a purely inductive branch. Assume that the circuit is operated at resonant frequency. Then:

$$X_L = X_C$$

Moreover, by substitution of X_L for R in Ohm's Law:

$$I_L = E/X_L$$

and by substitution of X_C for R;

$$I_C = E/X_C$$

so that:

$$I_L = I_C$$

where I_L and I_C represent current magnitudes. Since the current in the inductive branch lags the applied voltage by 90° and since the current in the capacitive branch leads the applied voltage by 90°, the two currents are 180° out of phase. Thus, since they are equal in magnitude, their sum is at every instant equal to zero. The parallel-resonant circuit is, thus, equivalent to an infinite resistance in series with the source.

For the purpose of the example, zero resistance was assumed for both branches of the circuit. In a practical circuit of this kind neither branch has zero resistance. Thus, in both branches of the circuit, the phase angle is slightly less than 90°. This has the effect of decreasing the effective resistance of the circuit, since the currents in the two branches are not exactly 180° out of phase with each other and so do not completely cancel each other.

Since in a resonant circuit the phase displacements of currents are equal and opposite, they cancel out each other, and the resultant current is in phase with the voltage. The circuit therefore acts as a high value of resistance. The lower the losses in the circuit, the higher this effective resistance becomes. Note that this effect is just opposite to that of the series circuit operated at resonance.

In a parallel circuit when X_C is greater than X_L, that is when the circuit is operated at a frequency lower than the resonant frequency, I_L (the current through the inductive branch) is greater in value than I_C (the current through the capacitive branch). The resultant current (I_T) lags the applied voltage; the circuit therefore acts as an inductance under the above circumstances.

On the other hand, when X_L is greater than X_C in a parallel L-C circuit; that is, when the circuit is operated at a frequency higher than the resonant frequency; the circuit acts as a capacitance.

3.9.5 Generalization of Ohm's Law

The impedance of a circuit may be expressed by the equation:

$$Z = E/I \qquad (25)$$

where:

Z = impedance in ohms

E = applied emf in volts

I = current in amperes

Here, emf and current are usually written in terms of effective (rms) values.

Since impedance in terms of resistance and reactance is given by:

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \qquad (23)$$

it follows that Ohm's Law, equation (1), is merely the special case of equation (23) characterized by $(X_L - X_C) = 0$. This is effectively the case in direct current circuits. Since any capacitance offers infinite impedance to d-c, no capacitance can be tolerated in any series portion of a d-c circuit. Moreover, capacitance in parallel can be ignored for the same reason. Inductance, on the other hand, offers zero reactance to d-c. Thus, for a d-c circuit:

$$Z = \sqrt{R^2} = R$$

3.10 POWER IN A-C CIRCUITS

As already noted, a purely inductive load or a purely reactive load in series with a source of emf does not consume any power since the power taken from the source is all returned to the source. These, of course, are theoretical cases. However, it can be shown that for a-c circuits:

$$P = EI \cos \theta \qquad (26)$$

where

P = power dissipated

E = effective value of emf

I = effective value of current

θ = angle between applied emf and current

Cos θ is called the power factor, because it must be multiplied by EI to give the true power dissipated.

For the case of the current in phase with the applied emf (i.e. $\theta = 0$), $\cos \theta = 1$ so that $P = EI$.

Components of a-c circuits are often rated in terms of volt-amperes (VI) rather than in terms of power. Here, V is the effective voltage drop through the load while I is the effective current through the load.

3.11 NON-SINUSOIDAL WAVEFORMS

In the discussion of alternating current up to this point, the form of the current has been assumed to be sinusoidal as shown in Figure 3-29.

In many specialized applications, however, different and complex wave shapes are used. An example is the square wave which is encountered in computing circuits (see Figure 3-43).

This is square rather than sinusoidal in shape, but it does repeat its square form over and over again indefinitely at intervals of T seconds. Thus the first positive alternation shown occupies the Time T/2; the first complete cycle occupies the Time T and so on.

Non-sinusoidal, periodic waves can be represented by a special set of sine waves integrally related to one another

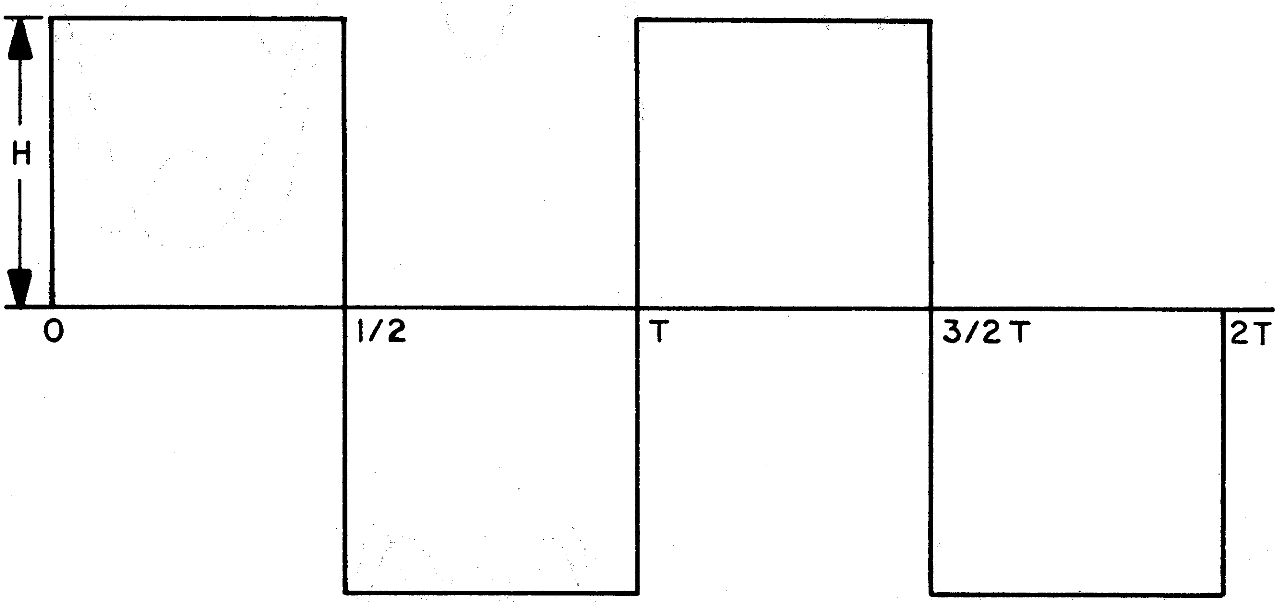


Figure 3-43

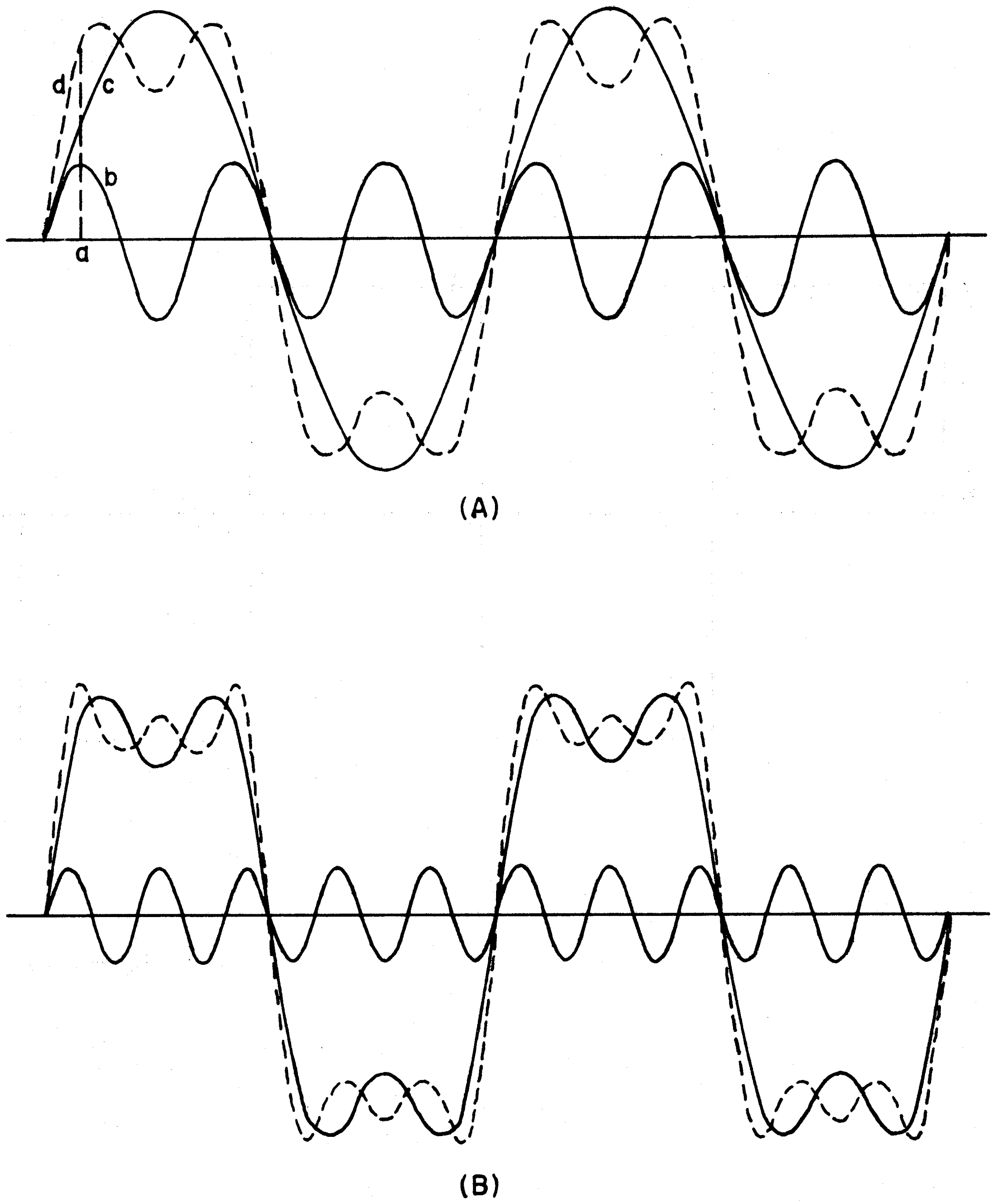


Figure 3-44

in frequency. The sinusoidal current flow for each sine wave component of a complex impressed voltage can be readily formulated; the total current is the sum of these sine wave current components, and represents the effect of the total complex impressed voltage.

It can be shown mathematically that the square wave of Figure 3-43 is composed of an infinite number of sine wave components of different frequencies. The sinusoidal component of lowest frequency is called the fundamental or first harmonic. The next component that occurs in the square wave is one of three times the frequency; it is called the third harmonic. Following this there is the fifth harmonic, then the seventh and so on; in short, all harmonics that are odd multiples of the fundamental.

The square wave is therefore found to have only odd harmonics. Other wave shapes may contain even harmonics as well. Each wave shape depends, however, not only on the number of harmonics present, but also upon their amplitudes and relative Timing (moment at which they cross the time axis).

For the square wave, the harmonics all pass through zero in a positive direction at the same time as the fundamental. The manner in which the harmonics combine to form the square wave is illustrated in Figure 3-44. In (A) the instantaneous values of the fundamental and third harmonics are added together to give wave shape shown in dotted lines. For example, at instant a the fundamental has the amplitude ac, and the third, harmonic has the amplitude ab. The sum of ac and ab is ad. This is the instantaneous amplitude of the resultant wave.

The latter, it will be observed, is steeper along the sides and flatter on top. The sum of the first two components is beginning to approach a square wave in shape. In (B) of Figure 3-44, the fifth harmonic is added to the resultant of (A), and gives rise to the second resultant shown in dotted lines in (B).

It will be noted that the new resultant is closer to a square wave in shape, and has many more ripples on top, but of smaller amplitude. As more and more harmonics are added, the sides approach the perpendicular; the top, a horizontal position and the ripples on the top become smaller and smaller as well as more numerous. Finally, when the infinite number of harmonics have been added, the wave becomes truly square in appearance.

Circuits which are designed to pass a particular non-sinusoidal waveform, should have a flat amplitude versus frequency characteristic for the sine wave frequencies composing that waveform. This follows from the fact that the wave shape depends upon the relative amplitudes of the individual components as stated above. Since a true square wave is composed of a fundamental and an infinite number of harmonics, it would be necessary to design a circuit with a flat response to an infinite number of frequencies in order to pass such a wave. This is not possible. However, circuits which can pass very good approximations of square waves can be de-

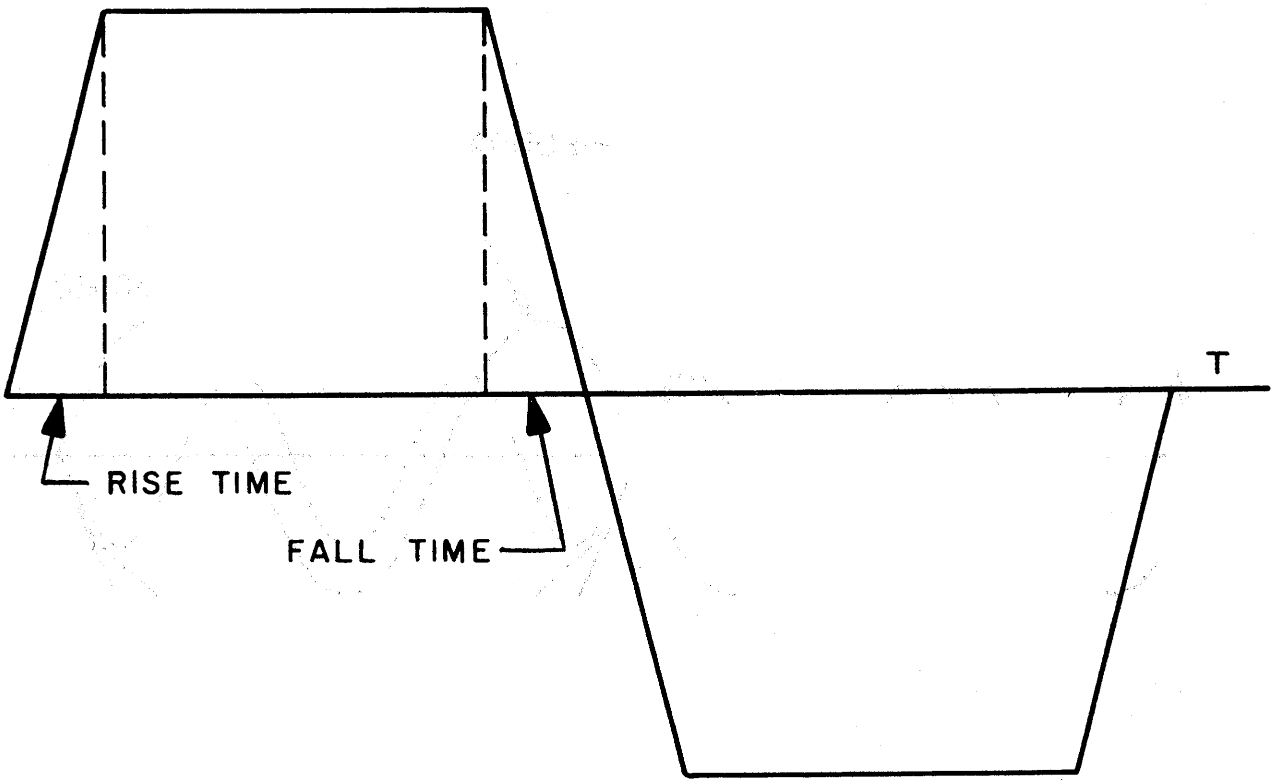


Figure 3-45

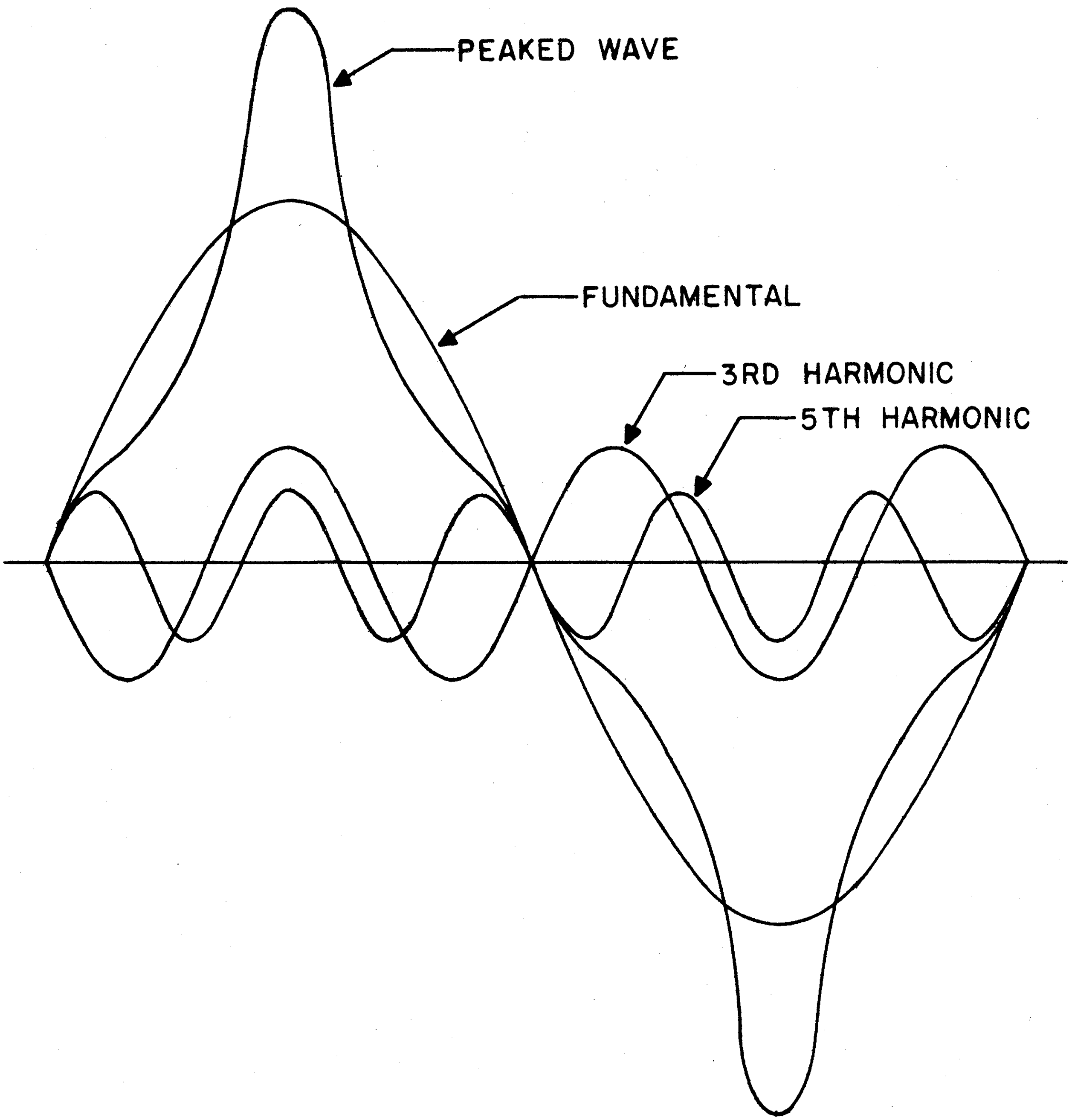
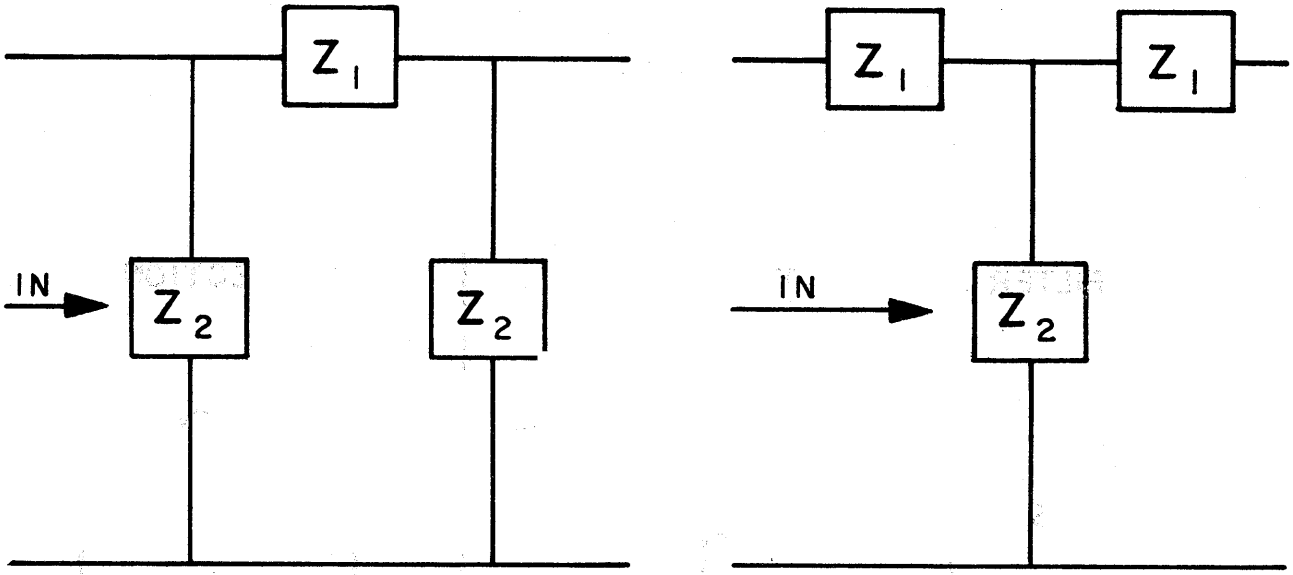
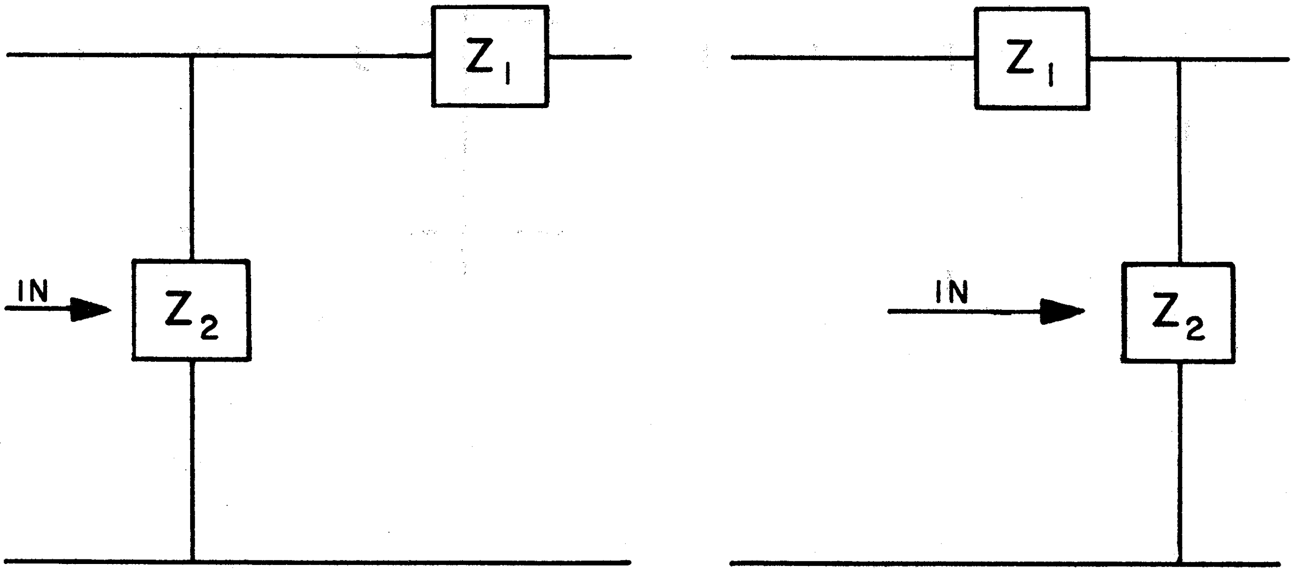


Figure 3-46



π SECTION

T SECTION



HALF π SECTION

HALF T SECTION

Figure 3-47

signed and find application in computer work. An approximate square wave is shown in Figure 3-45. The time required for the wave to rise from 10% of maximum to 90% of maximum is called the rise time. The time required to fall from 90% of maximum to 10% of maximum is called the fall time. This definition holds regardless of whether the maximum is positive or negative.

The square wave is just one of many special waveforms which may be encountered in computer work. Another common example is the peaked wave or spike shown in Figure 3-46. This, like the square wave, is composed of a fundamental and odd harmonics. However, as can be seen, the different time relationships between the fundamental and harmonics result in a waveform of a very different character.

3.12 FILTERS AND DELAY NETWORKS

Filters and delay networks are made up of resistances, inductances, capacitances, and mutual inductances or a combination of these parameters hooked up in configurations that are required by the function of a particular network. The most common configurations are the T sections and the π sections shown in Figure 3-47. These may be divided into half T sections or half π sections, as shown.

3.12.1 Filters

The description of filters in this section deals with their ability to either pass or block different frequencies. That is, a filter is a four terminal network which transmits fully a continuous band of frequencies and introduces attenuation to all other frequencies, or conversely, a band eliminater filter may attenuate a continuous band of frequencies and pass all others. Filters are classified as high pass, low pass, band pass, and band elimination, according to the band of frequencies to be passed or eliminated. Since no filter is perfect, a practical filter will offer some attenuation to frequencies in the pass band. Neither is infinite impedance present for frequencies outside the pass band. Moreover, attenuation does not change abruptly from minimum to maximum. That is, a small frequency range is required to change from minimum to maximum attenuation.

It must also be remembered that filters are part of coupling circuits and hence are designed to operate between a given load impedance and a given generator impedance. Thus, the design of a filter circuit is governed by these factors. In general, a low pass filter is obtained by letting Z_1 be inductive and Z_2 be capacitive in the equivalent π and T networks of Figure 3-47. Since the impedance of an induction rises with higher frequencies, the impedance of a low pass filter will vary directly with the input frequency. On the other hand, since the capacitive reactance decreases with an increase in frequency, a high band pass filter would have Z_1 capacitive and Z_2 inductive. For a band elimination filter on the other hand, each Z_1 block might be a parallel L-C combination, while each Z_2 block might be a series L-C combination. These L-C combinations would be selected so as to be resonant at the center of the band to be eliminated.

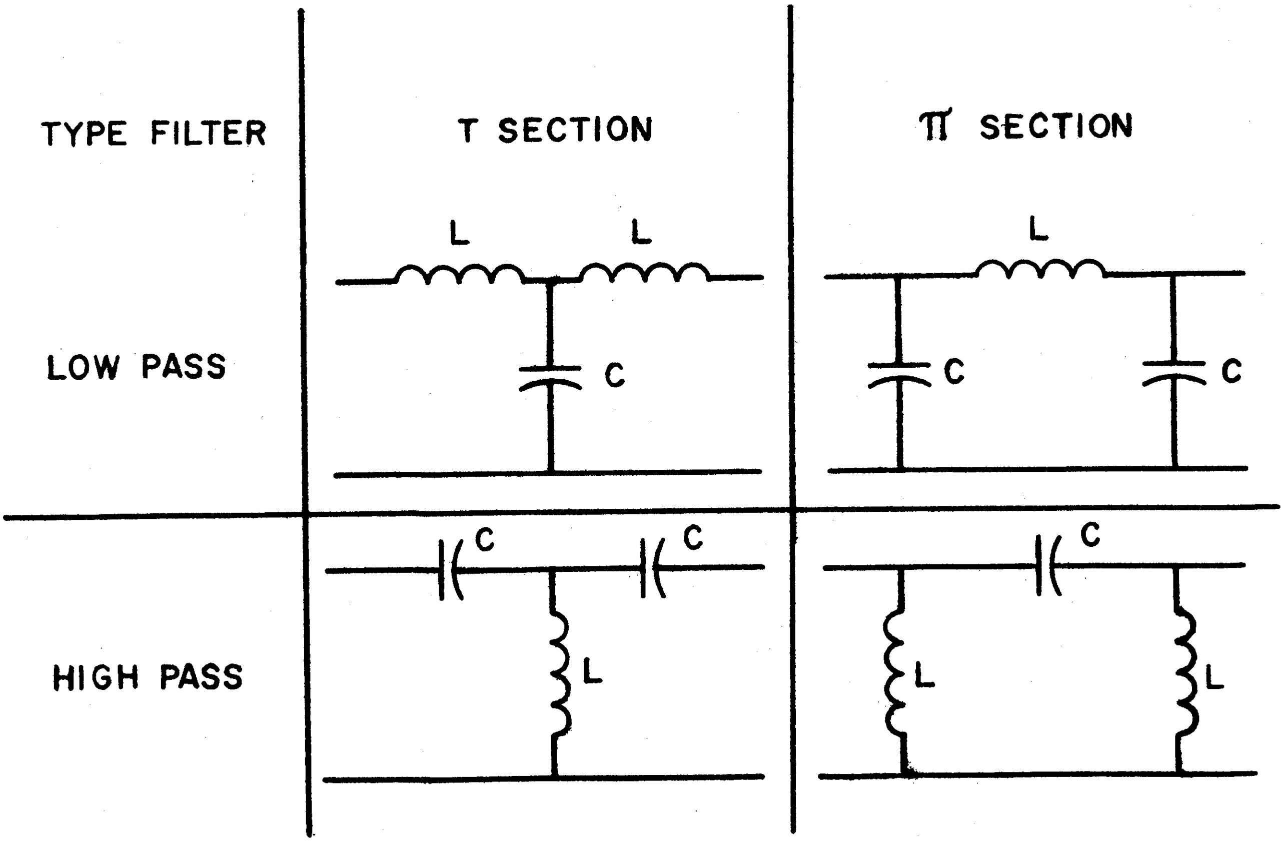

TYPE FILTER
T SECTION
π SECTION
LOW PASS
HIGH PASS
L
L
C
L
C
C
C
C
C
L
L
L

Figure 3-48

Figure 3-48 illustrates T section and π section filters with proper parameter arrangements.

3.12.2 Delay Networks

Filter networks such as those already described have, in addition to an attenuation factor, a propagation constant. This may be stated in terms of a single section of a complex filter. In studying capacitive and inductive circuits, it should be understood that a phase shift takes place in each section of a filter. Thus, a signal passing down a delay line, which is usually composed of several filter sections, has its phase changed at each section. And since phase is related to time, a delay line can be an effective means of retarding the arrival time of a signal. In computers, delays obtainable from capacitor and inductor networks range from a fraction of a microsecond to a few microseconds. The attenuating effect of electrical delay lines in computers requires amplification and reshaping of signals that pass through such delay lines.

3.13 TIME CONSTANTS

If a battery or other d-c source is applied to a series R-C load, current flows until the potential difference across the capacitance is equal to the emf of the source. The time required for the potential difference to rise to a value equal to 63% of the applied emf is called the time constant of the circuit and is given by the following equation:

$$T = RC$$

where:

T = time constant in seconds
R = resistance in ohms
C = capacitance in farads

Similarly, if a battery or other d-c source is applied to a series R-L load, the current increases at a rate determined by the magnitude of the emf and the inductance of the load until it reaches a maximum value determined by the resistance of the load. The time required for the current to reach 63% of its maximum value is called the time constant of the circuit and is given by the equation:

$$T = \frac{L}{R}$$

where:

T = time constant in seconds

R = resistance in ohms

L = inductance in henries

For example, if an R-C circuit, R = 500 ohms and C = 6 microfarads then:

$$t = R \times C = 500 \times 6 \times 10^{-6} = .003 \text{ second}$$

For an R-L circuit, where the inductance is .005 henry, and the resistance is 500 ohms, the time constant is

$$t = \frac{L}{R} = \frac{.005}{500} = .00001 \text{ second or 10 microseconds}$$

3.14 GENERAL NETWORK LAWS

A network may be described as a circuit possessing a combination of resistance, capacity and inductance, and in general is a combination of series and parallel arrangements. A number of theories and laws have been developed to determine the behavior of networks, and their application is essentially simple.

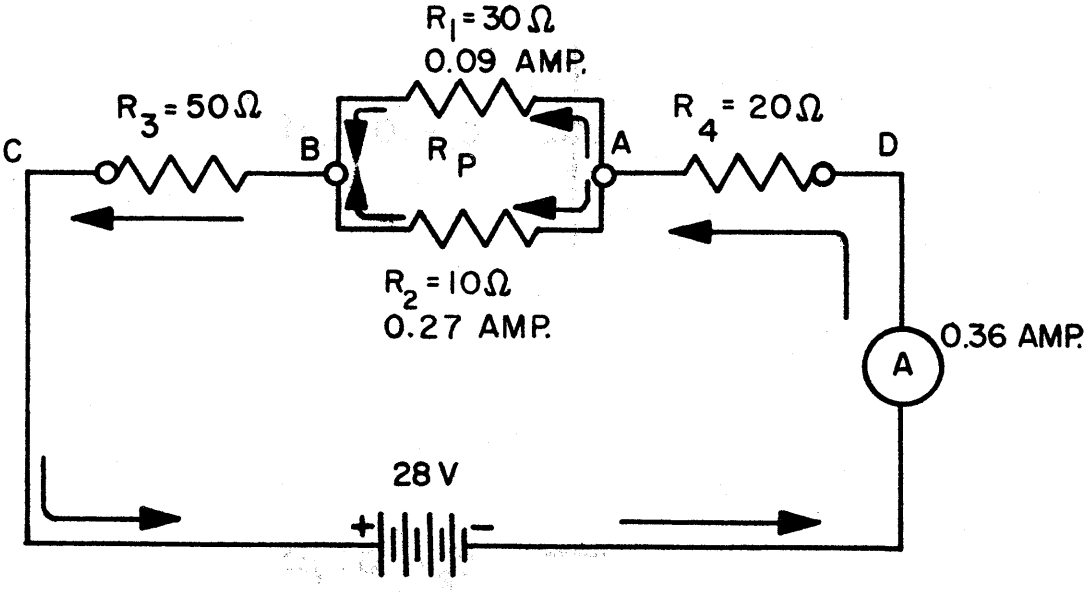


Figure 3-49

3.14.1 Kirchhoff's Current Law

This law states that in any network, the total current leaving a given point must be exactly equal to the current approaching that point. If the convention is adopted that current approaching a point has one sign and current leaving a point has a different sign, then this law can be stated as follows: The algebraic sum of all currents at any point in a network is zero.

In Figure 3-49 the current flow is toward D from the negative battery terminal as the arrows indicate. The current from A to B through R_1 is calculated to be .09 ampere, and the current through R_2 is .27 ampere. The current leaving point B is therefore .36 ampere as determined by the sum of the two currents and ammeter A. If the currents toward B are considered positive and the current away from B negative, then the algebraic sum at B is $.09 + .27 - .36 = 0$ as defined by Kirchhoff's Law.

3.14.2 Kirchhoff's Voltage Law

Kirchhoff's Voltage Law states that in a network, the sum of the voltage drops around any closed path equals the emf applied to the given path. This is illustrated by Figure 3-49. By equation (10)

$$R_p = \frac{R_1 R_2}{R_1 + R_2} = \frac{30 \times 10}{30 + 10} = \frac{300}{40} = 7.5 \text{ ohms}$$

where R_p is the resistance of the parallel combination. Then the voltage drop is: for R_p, $E = IR = .36 \times 7.5 = 2.7$ volts;

for R_3, E = IR = .36 x 50 = 18 volts; and R_4, E = IR = .36 x 20 = 7.2 volts. The sum of these individual drops is 2.7 + 7.2 = 27.9 volts which is approximately the applied voltage (the slight difference is due to the percentage of error introduced going beyond one decimal place).

3.15 CAPACITIVE VOLTAGE DIVIDERS

Capacitors can be used as alternating current voltage dividers. If a set of n capacitors is connected in series, then the equation for the voltage drop, V_1, across capacitor C_1 is:

$$V_1 = E \frac{X_{C_1}}{X_{C_1} + X_{C_2} + \ldots + X_{C_n}} \tag{26}$$

Similarly, the equation for the voltage drop, V_2, across capacitor C_2 is:

$$V_2 = E \frac{X_{C_2}}{X_{C_1} + X_{C_2} + \ldots + X_{C_n}}$$

and so on, where:

E = emf applied across the entire set

$X_{C_1}, X_{C_2}, \ldots, X_{C_n}$ are the capacitive reactances of the individual capacitors.

Since capacitive reactance is inversely proportional to capacitance, the reciprocal of capacitance can be used in the above equations rather than capacitive reactance. Thus, for example, equation (26) becomes:

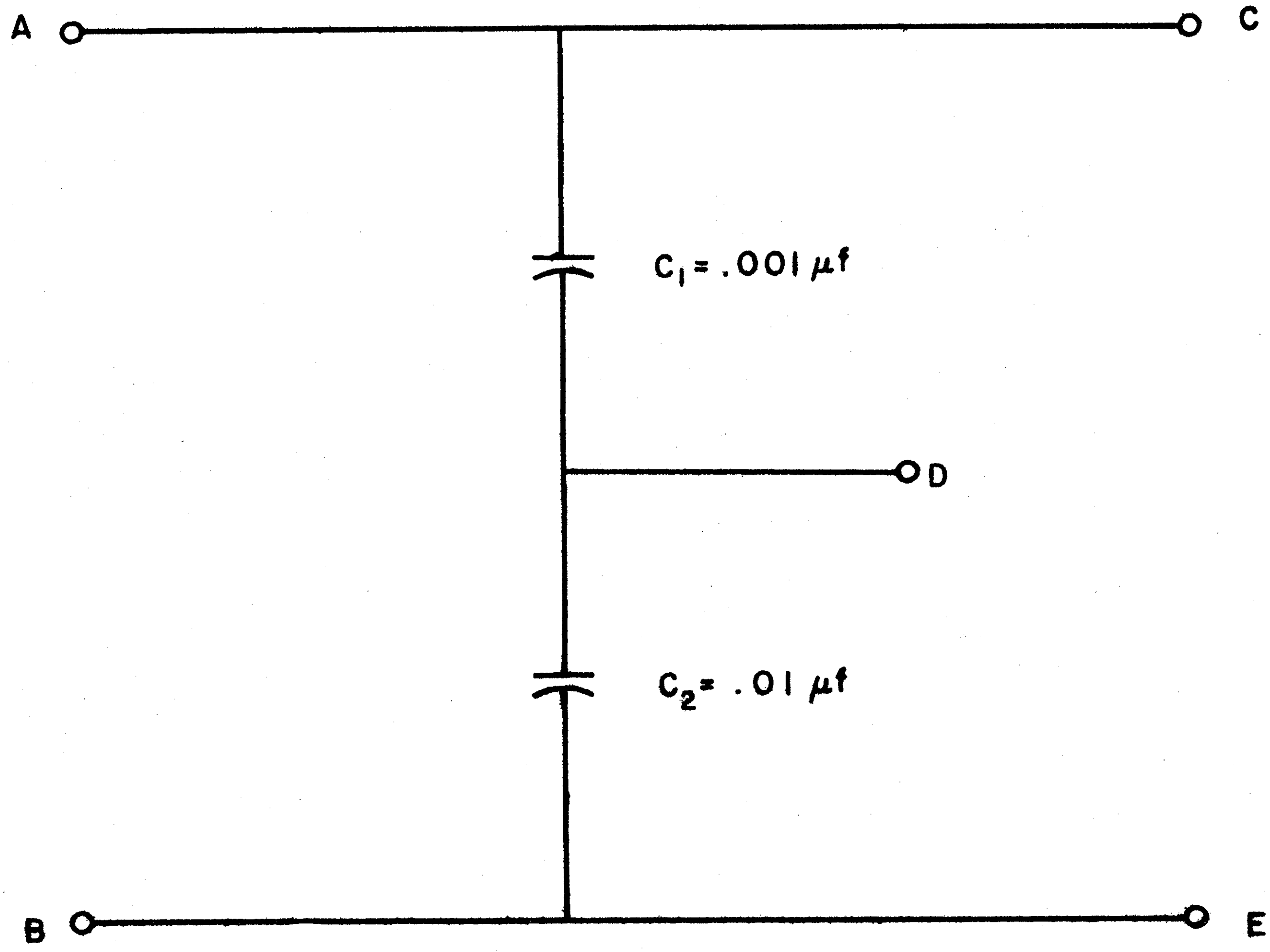


Figure 3-50

$$V_1 = E \frac{\frac{1}{C_1}}{\frac{1}{C_1} + \frac{1}{C_2} + \ldots + \frac{1}{C_n}} \qquad (27)$$

To illustrate, consider Figure 3-50. Here, capacitor C_1 = .001 microfarad and capacitor C_2 = .01 microfarad. Substituting these values into equation (27) yields:

$$V_1 = E \left(\frac{\frac{1}{.001}}{\frac{1}{.001} + \frac{1}{.01}} \right)$$

$$= E \left(\frac{1}{1 + 10} \right)$$

$$= \left(\frac{1}{11} \right) E$$

$$V_2 = E \left(\frac{\frac{1}{.01}}{\frac{1}{.01} + \frac{1}{.001}} \right)$$

$$= E \left(\frac{1}{1 + 1} \right)$$

$$= \left(\frac{10}{11} \right) E$$

To provide for the discharge of a capacitor once the applied voltage has been removed, it is standard practice to shunt each capacitor with a high resistance. These resistances should be in the same ratio as the voltage ratios determined by the capacitors alone. This is to avoid upsetting the voltage distribution.

3.16 IMPEDANCE MATCHING

The maximum amount of power is transferred from a generator to a load when:

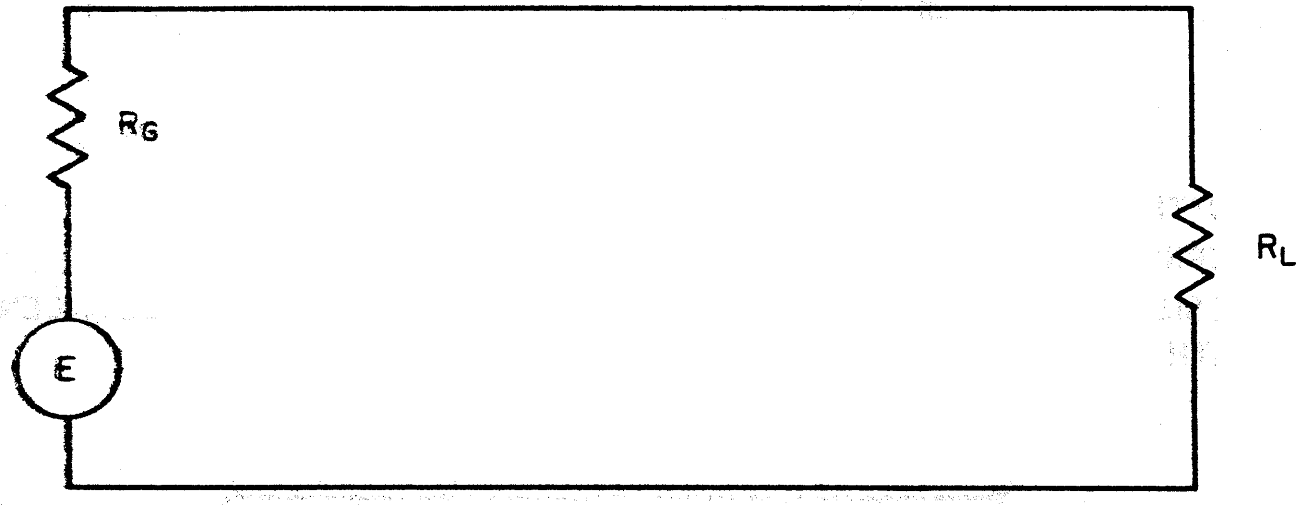


Figure 3-51

$$R_G = R_L$$

where:

R_G = internal resistance of generator

R_L = resistance of load

In order to understand this relationship, consider the simple circuit illustrated in Figure 3-51. Since the circuit is purely resistive, the current, I, is in phase with the applied emf, E, so that the power dissipated by the load P_L is given by the following equation:

$$P_L = V_L I \qquad (28)$$

where:

V_L = voltage drop through the load

By equation (3),

$$V_L = \frac{R_L}{R_G + R_L} E \qquad (29)$$

By Ohm's Law,

$$I = \frac{E}{R_L + R_G} \qquad (30)$$

Thus, assuming E and R_G to be constant, an increase in R_L decreases the current as can be seen by an inspection of equation (30), but, at the same time, increases the proportion of the emf that appears across the load as can be seen by an inspection of equation (29). For R_G greater than R_L, the increase in the proportion of emf appearing across the load per unit change of R_L is greater than the decrease in current per unit change of R_L. Thus, an increase of R_L results in an increase of power dissipated through the load. On the other hand, for R_G less than R_L, the increase in the proportion of emf appearing across the load per unit change of R_L is less than the decrease in current per unit change of R_L. Thus, an increase of R_L results in a decrease of power dissipated through the load. From this, it follows that the maximum power transfer occurs when $R_G = R_L$.

In order to clarify this relationship consider the following three cases:

a. $R_G = R$, $R_L = 1/2\text{-}R$ (i.e. $R_G > R_L$)

Substituting these values into equations (29) and (30):

$$V_L = \frac{1/2\text{-}R}{R + 1/2\text{-}R} E = \frac{E}{3} \tag{31}$$

$$I = \frac{E}{1/2\text{-}R + R} = \frac{E}{3/2\text{-}R} \tag{32}$$

Substituting equations (31) and (32) into equation (28):

$$P_L = \frac{E^2}{9/2\text{-}R} = \frac{2E^2}{9R} = \frac{2}{9} \; \frac{E^2}{R}$$

b. $R_G = R$, $R_L = R$ (i.e. $R_G = R_L$)

Substituting these values into equations (29) and (30):

$$V_L = \frac{R}{2R} E = \frac{E}{2} \tag{33}$$

$$I = \frac{E}{2R} \tag{34}$$

Substituting equations (32) and (34) into equation (28):

$$P_L = \frac{E^2}{2R} \quad \frac{1}{2} \quad \frac{E^2}{R}$$

c. $R_G = R$, $R_L = 2R$ (i.e. $R_G < R_L$)

Substituting these values into equations (29) and (30):

$$V_L = \frac{2R}{R + 2R} E = \frac{2}{3} \tag{35}$$

$$I = \frac{E}{R + 2R} = \frac{E}{3R} \tag{36}$$

Substituting equations (35) and (36) into equation (28):

$$P_L = (\frac{2}{3}E)\ (\frac{E}{3R}) = \frac{2}{9}\ \frac{E^2}{R}$$

A comparison of the results obtained in these three cases reveals that two and one half times as much power is transferred when $R_G = R_L$ as when $R_G = 2R_L$ or when $R_L = 2R_G$.

In order to simplify this discussion, pure resistances have been considered. In the general case of impedance, the maximum power is transferred from a generator to a load when

$$Z_G = Z_L \text{ and } \theta_G = - \theta_L$$

where:

Z_G = internal impedance of generator

Z_L = impedance of load

θ_G = generator impedance angle

θ_L = load impedance angle

Thus, if the generator impedance is effectively a capacitance in series with a resistance, the load impedance should be effectively an inductance in series with a resistance. Moreover, the capacitive reactance of the generator should be of the same magnitude as the inductive reactance of the load, and the generator and load resistances should be equal.

3.17 TRANSMISSION LINES

A transmission line is a line which is used to conduct or guide electrical energy from one point to another. Its main requirement is to deliver as much power as possible from a source at one end of the line to a load at the other end of the line. In order to obtain maximum power transfer the

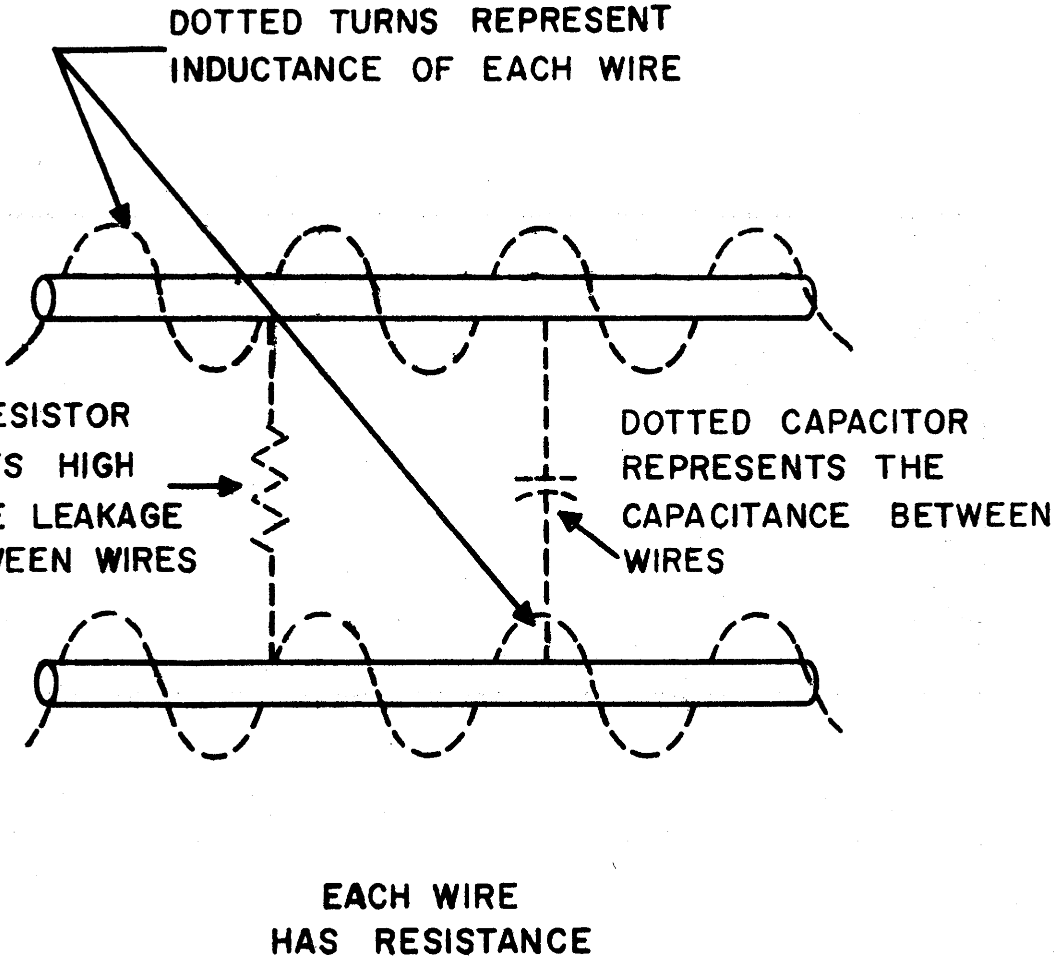

DOTTED TURNS REPRESENT INDUCTANCE OF EACH WIRE
DOTTED RESISTOR REPRESENTS HIGH RESISTANCE LEAKAGE PATH BETWEEN WIRES
DOTTED CAPACITOR REPRESENTS THE CAPACITANCE BETWEEN WIRES
EACH WIRE HAS RESISTANCE

Figure 3-52

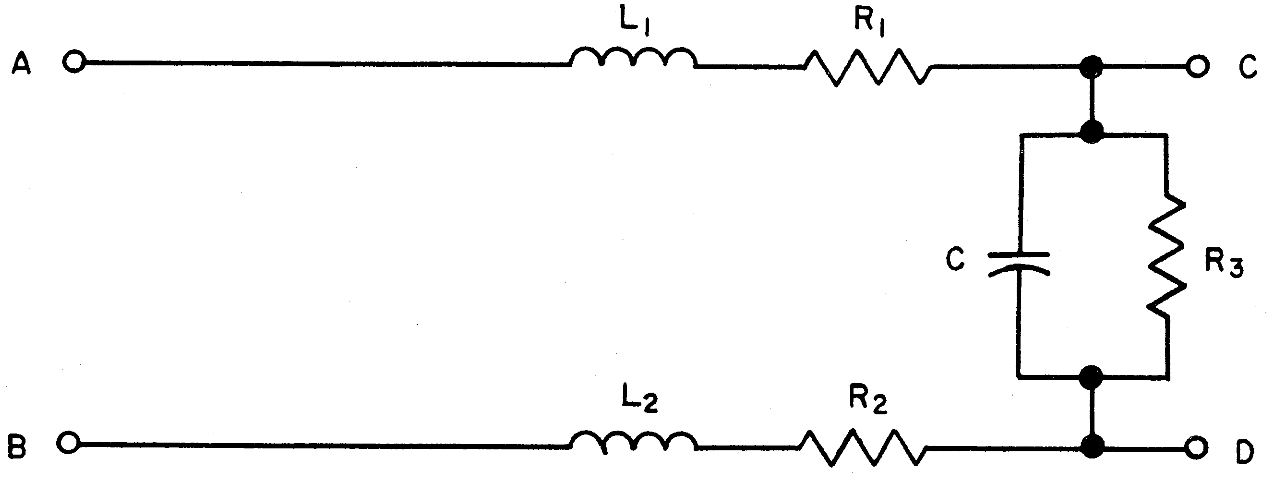


Figure 3-53

efficiency of the line must be high (nearly all the power entering the line at the sending end must reach the receiving end). The source impedance must be matched to the input impedance of the line and the load impedance must be matched to the output impedance of the line.

3.17.1 Transmission Line Impedance

The impedance of a transmission line depends on such properties as inductance in series with the line, capacitance across the line, resistance leakage paths across the line, and certain radiation losses, besides the resistance of the line itself. These properties are distributed throughout the entire line and cannot be separated from each other. Figure 3-52 shows these properties of resistance, inductance, capacitance, leakage resistance combined in a very short section of a two wire transmission line. This diagram is actually not a true representation of a transmission line because it shows the evenly distributed capacitance as a single lumped capacitor and the distributed leakage resistance as a lumped leakage path. However, if the section is very short compared to the total length of the line as measured in wave lengths, this approximation is close enough for practical purposes. Figure 3-53 represents all four properties lumped by conventional symbols. As a matter of fact, a transmission line may be thought of as an infinite number of circuits such as that shown in Figure 3-53 connected in cascade (i.e. connected so that the output terminals of the nth circuit are connected to the input terminals of the (n + 1)th circuit. This mode of analysis implies that each of the circuits is infinitely small. R_1 and R_2 represent the resistance of each wire, L_1 and L_2 the inductance of each wire, C the capacitance and R_3 the leakage resistance between the wires.

The properties discussed in the preceding paragraph are encountered for the following reasons: Any conductor has electrical resistance, depending on its length, cross-sectional area and the material it is composed of. When a current flows through this conductor magnetic lines of force are set up which encircle the wire. Upon the collapse of this field a certain amount of energy which is contained in the magnetic field is returned to the circuit to keep the current flowing a little longer. This is the property of inductance. When two of these conductors are placed near each other with air or any other insulating material between them acting as a dielectric, they behave like the plates of a capacitor. Many current leakage paths exist between these two conductors because no insulator, even air, is perfect and therefore a certain insulation resistance exists between the wires.

3.17.2 Characteristic Impedance

The characteristic impedance of a transmission line is the impedance across the input of a theoretically infinite line. Usually the characteristic impedance of a transmission line is given by the manufacturer who finds it as follows: The impedance of a single section such as shown in Figure 3-53 is determined. The impedance between terminals AB (Figure 3-53) can be calculated by the use of series-parallel impedance formulas, if the impedance across CD is known. However,

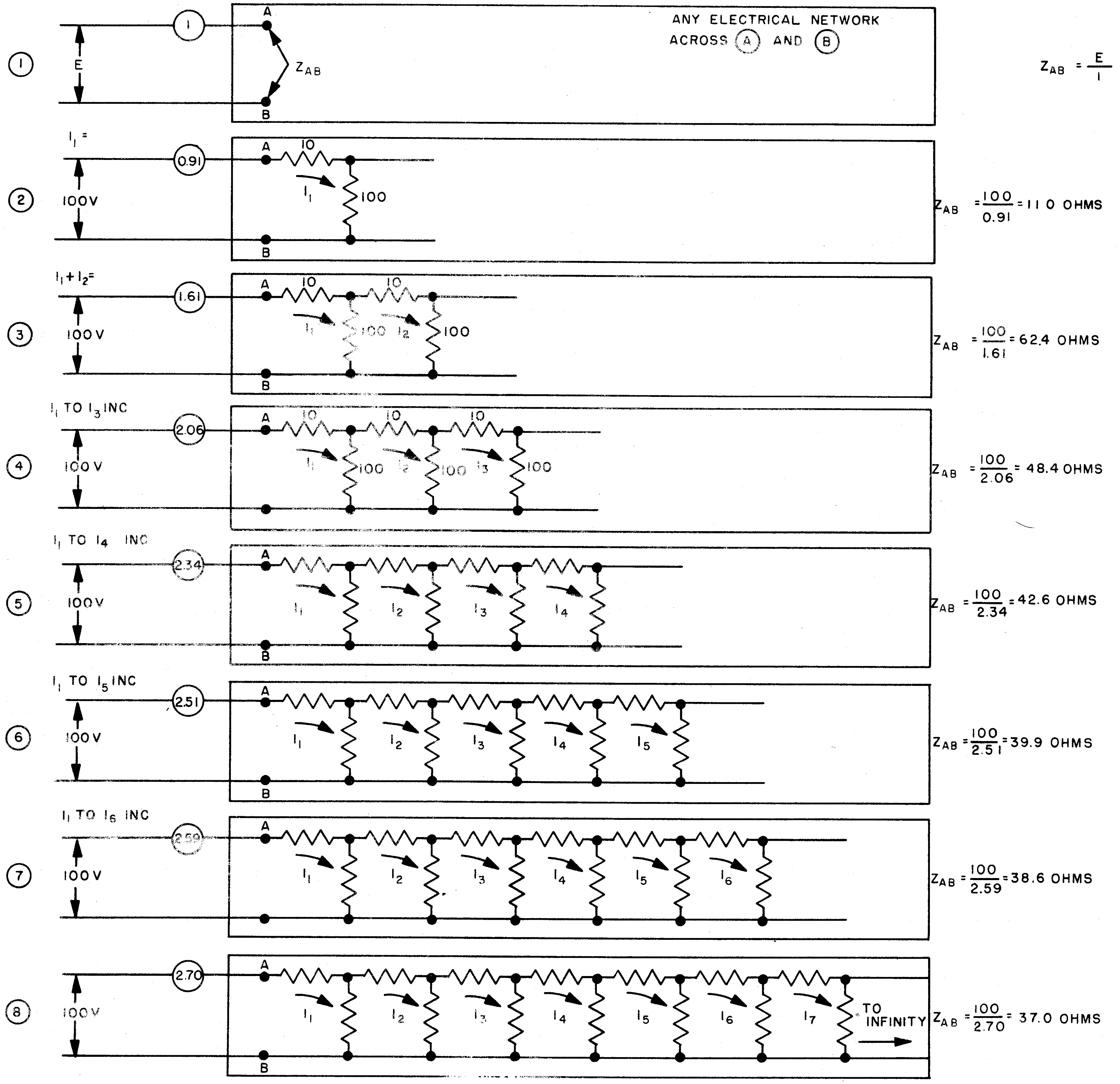

ANY ELECTRICAL NETWORK ACROSS A AND B
① $Z_{AB} = \frac{E}{I}$
② $I_1 =$ 0.91, 100V, $Z_{AB} = \frac{100}{0.91} = 110$ OHMS
③ $I_1 + I_2 =$ 1.61, 100 V, $Z_{AB} = \frac{100}{1.61} = 62.4$ OHMS
④ I_1 TO I_3 INC 2.06, 100 V, $Z_{AB} = \frac{100}{2.06} = 48.4$ OHMS
⑤ I_1 TO I_4 INC 2.34, 100V, $Z_{AB} = \frac{100}{2.34} = 42.6$ OHMS
⑥ I_1 TO I_5 INC 2.51, 100 V, $Z_{AB} = \frac{100}{2.51} = 39.9$ OHMS
⑦ I_1 TO I_6 INC 2.59, 100 V, $Z_{AB} = \frac{100}{2.59} = 38.6$ OHMS
⑧ 2.70, 100 V, TO INFINITY, $Z_{AB} = \frac{100}{2.70} = 37.0$ OHMS
A
B
E
I
Z_{AB}
10
100
I_1
I_2
I_3
I_4
I_5
I_6
I_7

Figure 3-54

since this section is only one small part of a longer line other similar sections are actually across terminals CD. If the impedance of each of these sections is known, the impedance at terminals AB can be calculated. With the addition of each of these sections the impedance at terminals AB has a new and lower value. However, after a multitude of these sections have been added, each successively added section has less and less effect on the impedance across AB. If sections are added to the line endlessly the line becomes infinitely long and a certain definite value of impedance is finally reached. If a load equal to the input impedance of the infinite line (i.e. characteristic impedance) is placed on the output end of any short or convenient length of line the same impedance appears at the input terminals of the line. This is the only value of impedance for any particular type and size of line which behaves in this manner. Thus the impedance of the source should be matched to this impedance in order to obtain maximum power transfer.

Figure 3-54 is a numerical example of the way small sections of line can be added successively to show the characteristic impedance. In 1 of Figure 3-54 a voltage (either direct or alternating) causes a current to flow and the ratio of E/I is equal to Z_{AB}, the impedance across terminals AB. It is noted as sections are added in Figure 3-54, 2 through 8, using a direct current of 100 v and an impedance composed of pure resistance that the impedance across AB is lowered and the current increases. However, the change in both becomes

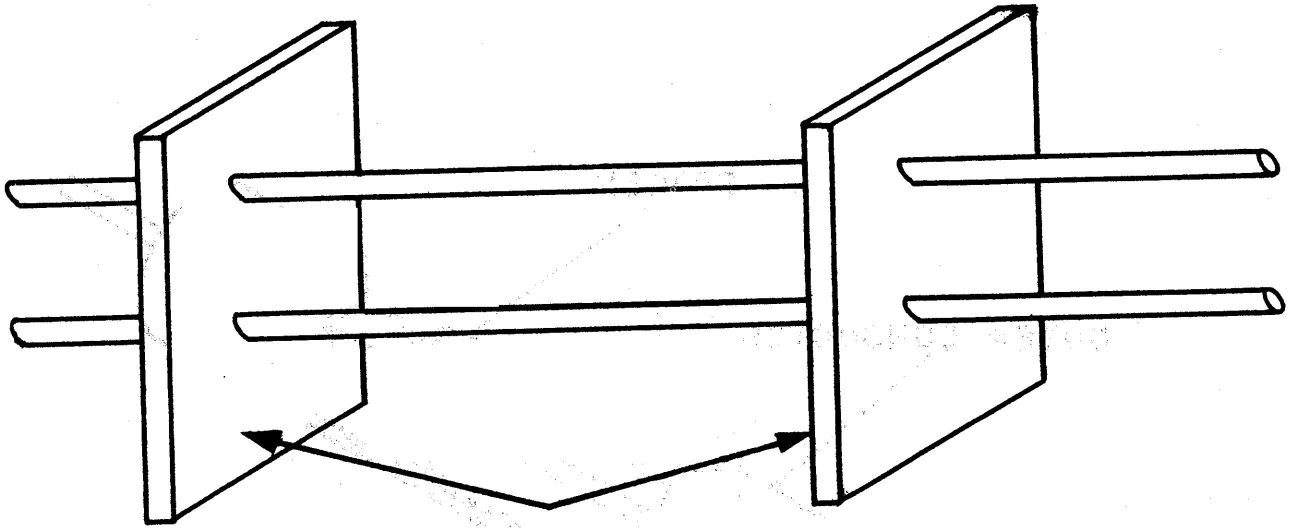


Figure 3-55

less and less as more sections are added. Finally, as an infinite number of sections are added the impedance across AB is reduced to 37 ohms and no further. This value of Z_{AB} then can be considered as Z_o, the characteristic impedance.

If 37 ohms is placed across the output of the first section instead of a second section, Z_{AB} will still be 37.0 ohms. The current will also have the same value as it had when the line had many sections. In fact if this 37.0 ohm resistor is connected across the output end of any of the networks of Figure 3-54 the input impedance Z_{AB} will be 37.0 ohms, the same as the input impedance of an infinite line. There is only one such value of impedance which a line assumes, if it is very long.

3.17.3 Types of Transmission Lines

There are four types of transmission lines. In general use they are the two-wire or parallel-conductor line, the concentric (coaxial) line, the twisted pair, and the shielded pair.

3.17.3.1 Parallel Two-Wire Line

The parallel two-wire line, Figure 3-55, is the most common type of transmission line. It consists of two parallel conductors which are maintained at a fixed distance by means of insulating spacers or spreaders at suitable intervals. Because of its ease of construction, economy, and efficiency this type is frequently used for commercial power lines, telephone and telegraph lines and as a connecting link between an antenna and transmitter or an antenna and receiver.

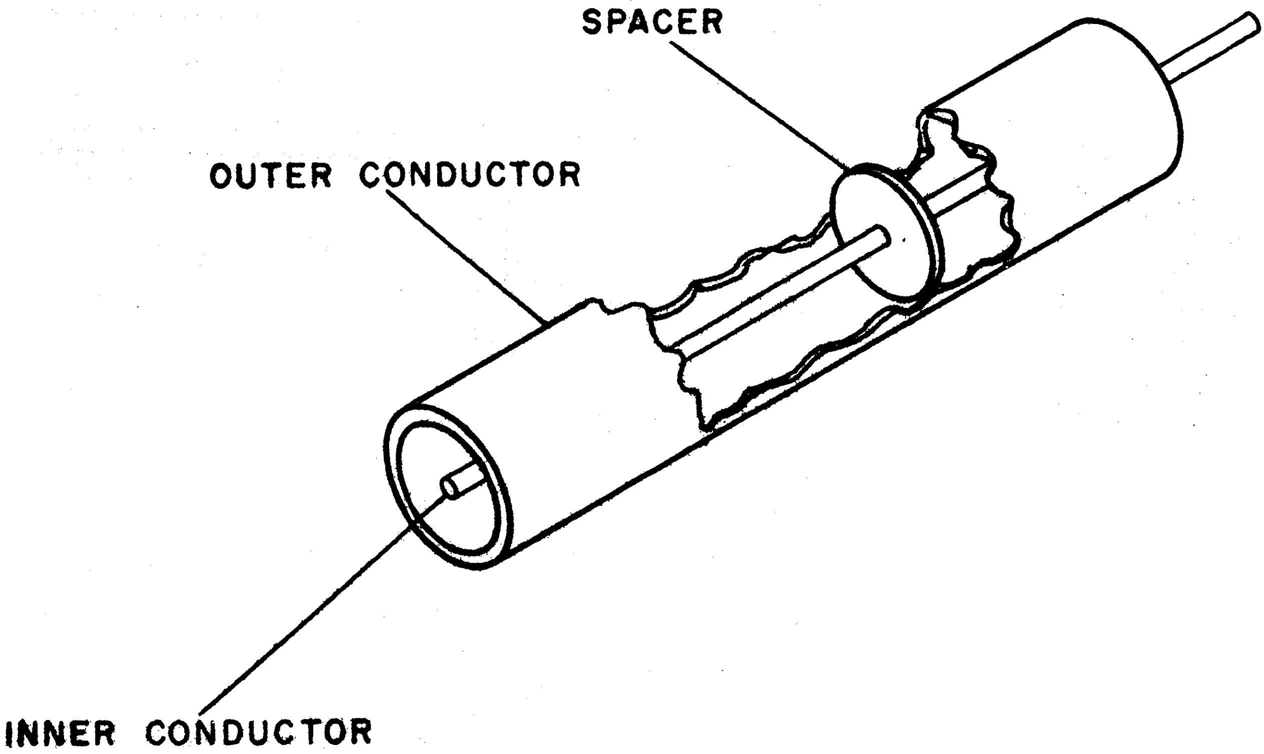


Figure 3-56

When used for 14-megacycles and lower frequencies the lines are generally spaced from 2 to 6 inches apart. A great disadvantage of the parallel-wire transmission line is that it has high radiation loss and therefore cannot be used in the vicinity of metallic objects at extremely high frequencies because of the greatly increased loss which results.

3.17.3.2 Concentric (Coaxial) Line

The concentric or coaxial line, Figure 3-56, is very efficient at extremely high frequencies. It is composed of a wire which is inside of and coaxial with a tubular outer conductor. The inner conductor may also be tubular in some cases. Insulating spacers or beads are placed at regular intervals to insulate the inner wire or conductor from the outer conductor. The spacers are made of pyrex, polystyrene, or some other material possessing good insulating qualities and low loss at high frequencies. Concentric cables are also made with the inner conductor consisting of flexible wire insulated from the outer conductor by a solid and continuous insulating material. When the outer conductor is made of metal braid flexibility may be gained but the losses are high.

In a coaxial line the radiation losses are kept down to a minimum, because no electric or magnetic fields extend outside of the outer conductor as is in the case with the two-wire parallel line. All the electric and magnetic fields that are present in the coaxial line exist in the space between the two conductors. Therefore the coaxial line is a perfectly shielded line.

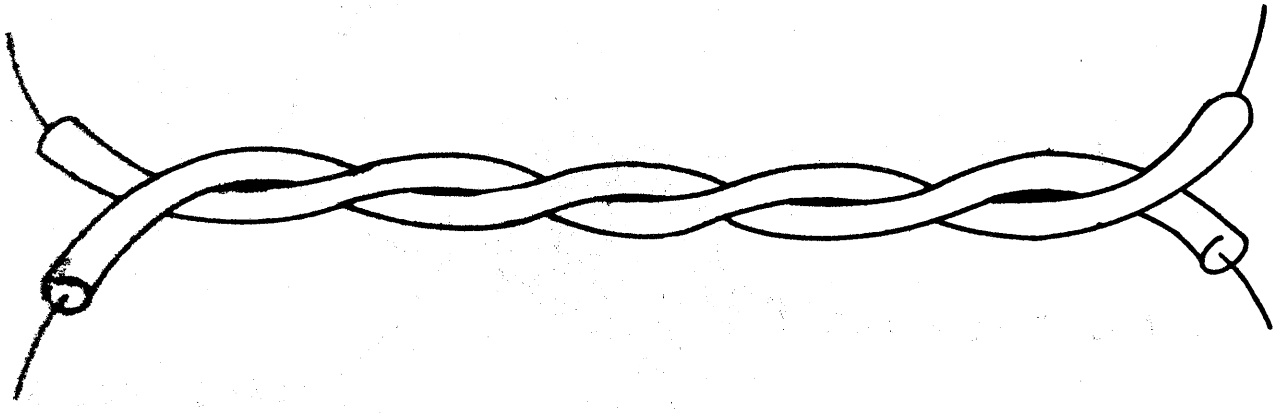


Figure 3-57

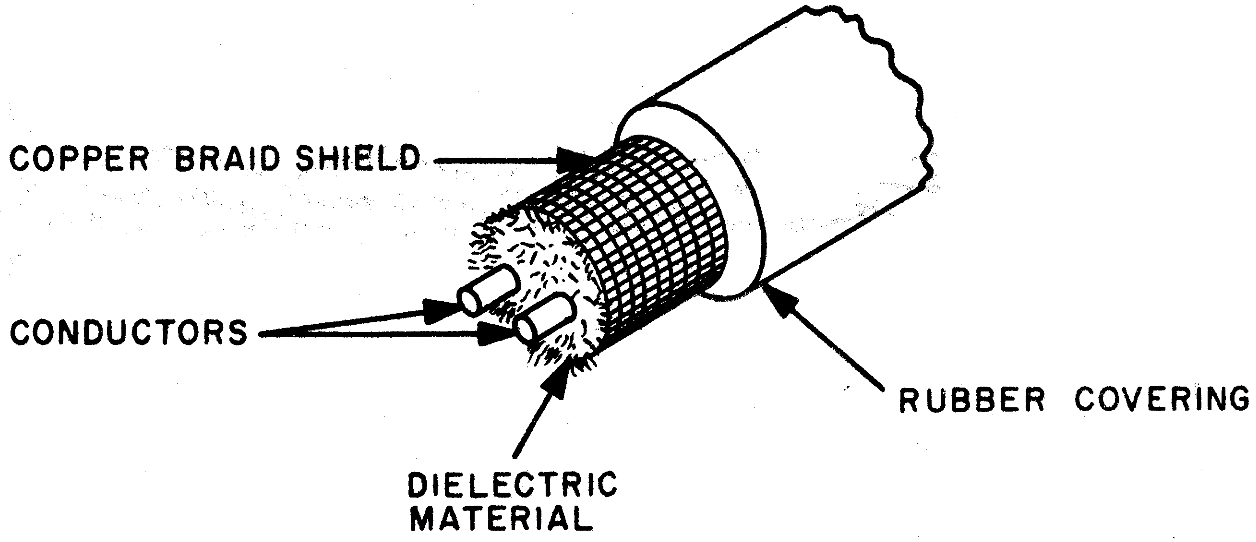

COPPER BRAID SHIELD
CONDUCTORS
RUBBER COVERING
DIELECTRIC MATERIAL

Figure 3-58

The advantages of an efficient concentric line are its very low radiation loss, due to the outer conductor acting as a shield for the inner conductor and the ease with which the line can be installed. Its disadvantages are that it is limited to short runs because of its loss at extremely high frequencies; it is comparatively more expensive for a given length of line; and it must be kept dry in order to prevent excessive leakage between conductors. In certain applications dry nitrogen at pressures ranging from 3 to 35 pounds per square inch may be used to fill in the lines in order to prevent the condensation of moisture. When first installed nitrogen is used to dry the line and a pressure is maintained to insure that leakage will be outward.

3.17.3.3 Twisted Pair Line

The twisted pair line, Figure 3-57, consists of two insulated wires twisted to form a flexible line without the use of spacers. It is generally used as an untuned line for low-frequency transmission. Due to the high losses occurring in the rubber insulation it is not used at the higher frequencies. Therefore the twisted pair line is mainly used for short distances at lower frequencies. Its main advantage is that it may be used where more efficient lines would not be feasible because of cost and mechanical considerations.

3.17.3.4 Shielded Pair Line

The shielded pair line, Figure 3-58, consists of two parallel conductors separated from each other and surrounded by an insulating dielectric material, such as the plastic, copaline. A copper-braid tubing surrounds the conductors and acts as a shield for them. A rubber or flexible composition coating covers this assembly to protect the line against moisture and friction. The shielded pair line has an outward appearance of an ordinary power cord for an electric motor.

The main advantange of the shielded pair line is that its two conductors are balanced to ground (the capacitance between each conductor and ground is uniform along the entire length of the line and the wires are shielded against pick-up of stray fields). This balance to ground is effected by the grounded shield which surrounds the conductors at a uniform spacing throughout their length. If this shielding is not present radiation can only be prevented if the current flow in each conductor is equal in amplitude thereby setting up equal and opposite magnetic fields which cancel out. However, this condition may only be obtained if the line is well in the clear of all obstructions, and the distance between the wires small. But if a line runs near some grounded or conducting surface, one of the conductors will be closer to the obstruction than the other. Depending upon the size of the obstruction a certain amount of capacitance will exist between the two conductors and the conducting surface over a portion of the line. A division of current flow will exist between each conductor, because of the parallel conducting path caused by the capacitance in each half of the line. There will be an inequality of current flow in the two conductors since one conductor is closer to the obstruction resulting in radiation due to the incomplete cancellation. The shielded line by maintaining balanced capacitances to ground eliminates such losses to a considerable degree.

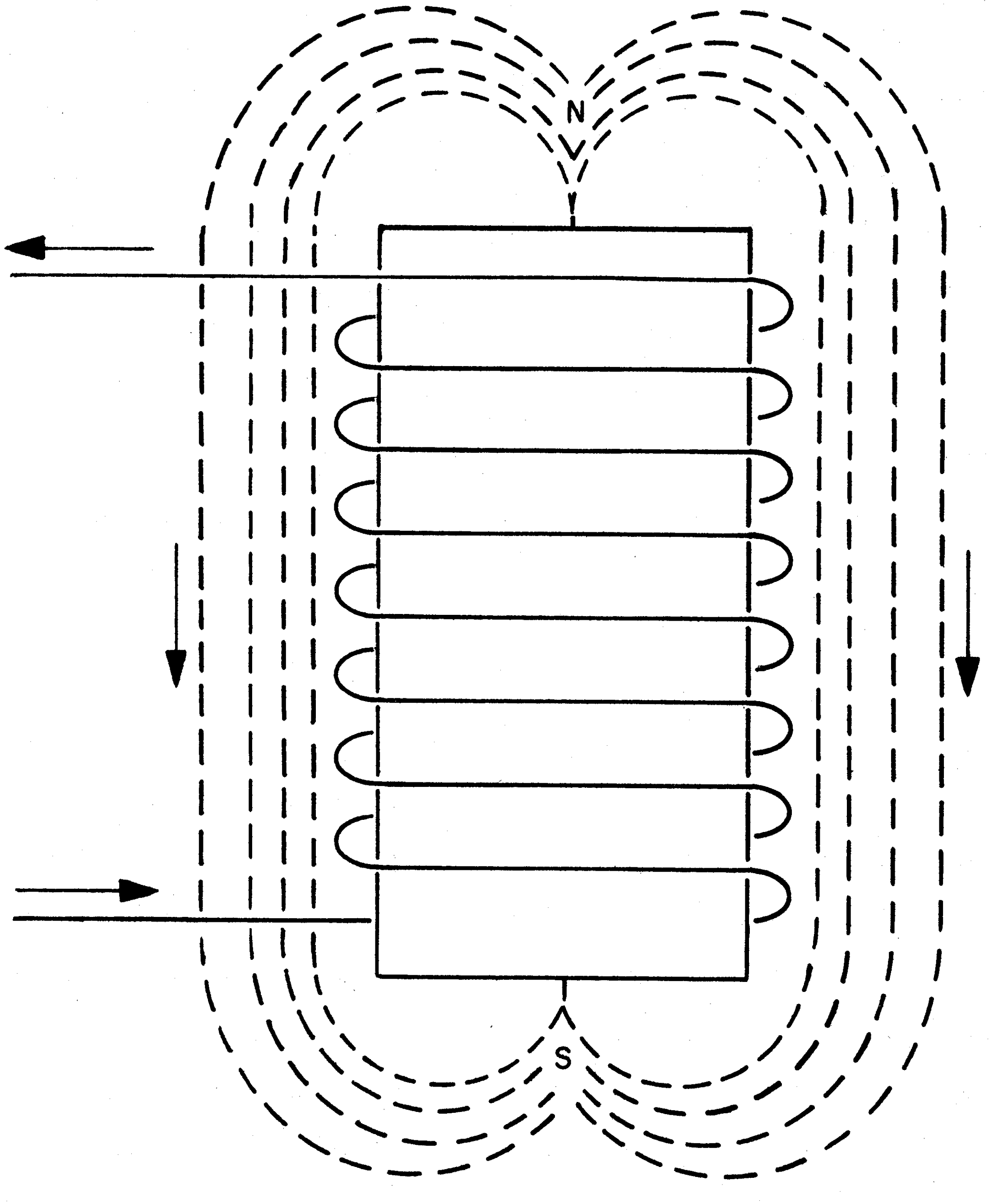


Figure 3-59

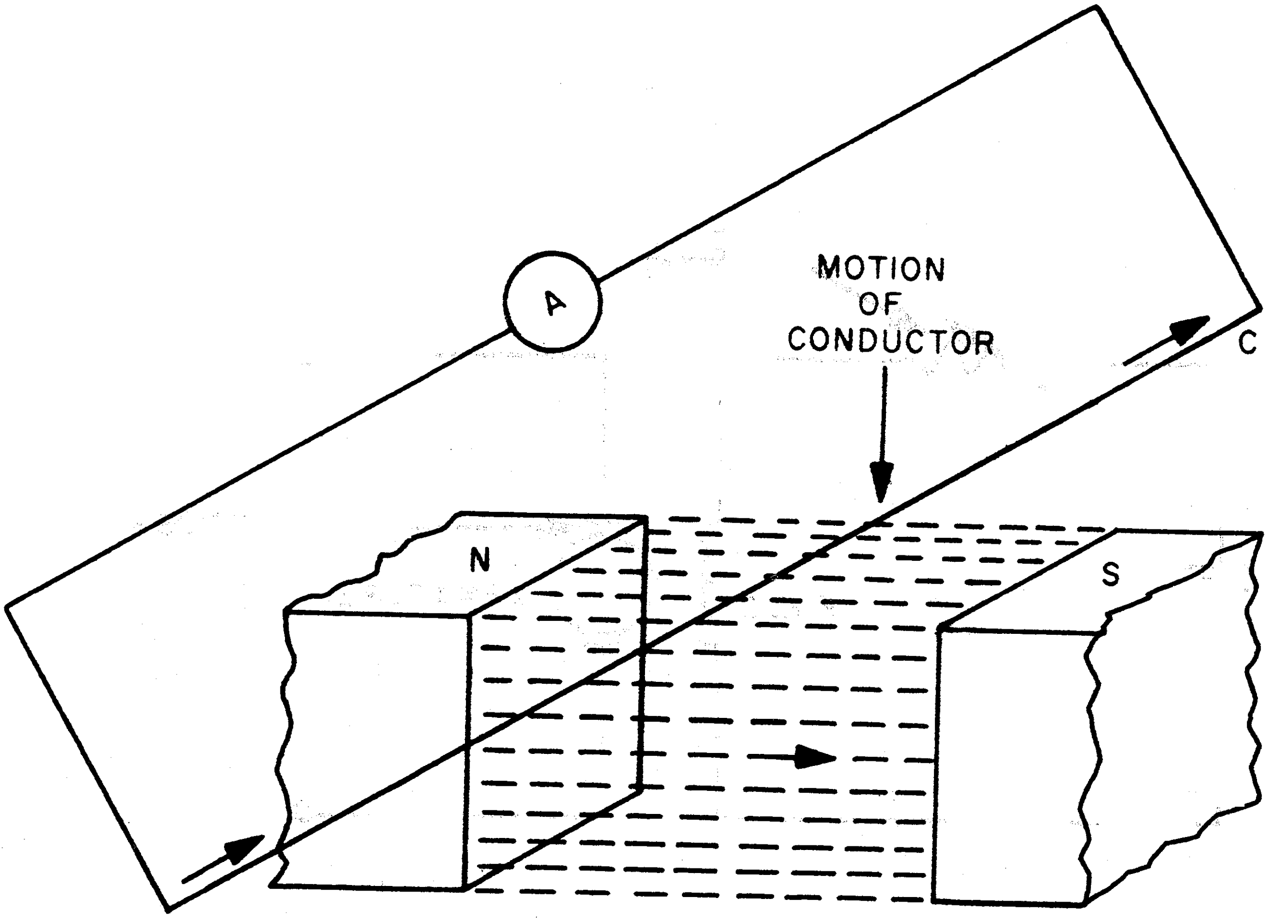


Figure 3-60

PART 3

CHAPTER 4

MOTORS AND GENERATORS

4.1 SOURCE OF ELECTRICITY

It has been shown in Chapter 2 on magnetism that there is a definite relationship between magnetism and electricity. Magnetism, as it has already been pointed out, is produced by electric currents. Any conductor carrying electricity has a magnetic field about it, and when a wire carrying a current is formed into a coil, the magnetic field passes through the coil and loops around the outside as shown in figure 3-59. The strength of this field depends upon the strength of the current flowing in the wire and the number of turns comprising the coil. When an iron core is placed within the coil the strength of the magnetic field is further increased.

A second relation that exists between magnetism and electricity, is the production of an electric voltage in a conductor moving across a magnetic field. This phenomenon is illustrated in Figure 3-60, which shows a voltage being induced in a conductor "C" as it is moved downward through the magnetic field in the air gap between the north and south poles of a magnet. This induced voltage causes an electron flow through the conductor in the direction of the arrows. The induced voltage can be increased either by a stronger magnetic field or by a more rapid motion of the conductor. The same voltage is produced if the conductor is kept stationary and the magnetic field moved upward across it.

There are many ways of obtaining electrical energy. For example, batteries convert chemical energy into electrical energy, photo cells convert light energy into electrical energy, thermocouples convert heat energy into electrical energy and piezo-electric crystals convert mechanical energy into electrical energy. However, this section is only concerned with the rotary sources such as the electric generator and associated equipment such as motors and controllers.

4.2 ELECTRIC GENERATOR

An electric generator is a device that converts mechanical energy into electrical energy. A generator comprises a conductor winding, or armature, which is caused to rotate in a magnetic field. Mechanical energy is expended in turning the armature, while electrical energy appears in the form of an emf induced in the armature as its individual turns cut across the lines of force of the magnetic field. A simple generator, in which the armature is a single loop of wire rotated in the field of a permanent magnet, is introduced in Section 3.2 of this Part. While this device illustrates the fundamental action of all generators, it differs in a number of ways from a practical generator.

The means of rotating the armature of the simple generator of Section 3.2 was not specified. In practical generators this function is usually provided by an electric motor. Thus, as

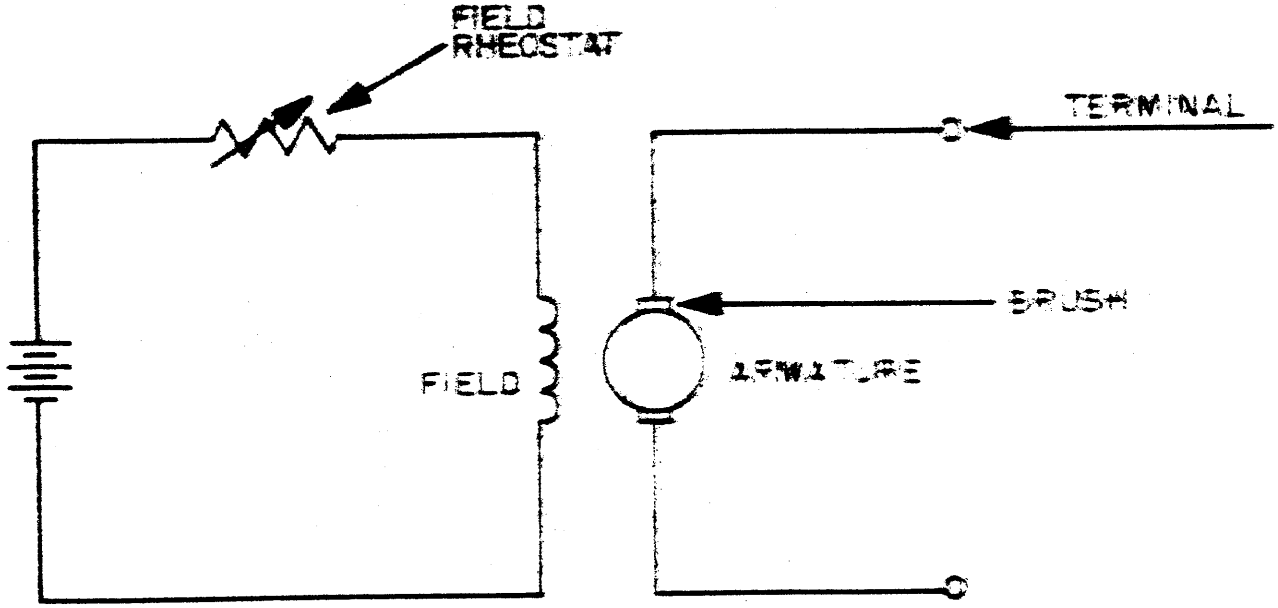


Figure 3-61

electrical energy is being generated by the one device, it is being dissipated by the other. However, the amount of electric power consumed by the motor which drives a generator armature is small compared to the amount of electric power supplied by the generator.

In actual practice, electromagnets take the place of the permanent magnet of Figure 3-30 to produce the magnetic field in which the armature rotates.

The field of an electric generator may either be separately excited or self-excited. Whenever current is supplied to the field from an external source, such as a battery it is called separately excited. Figure 3-61 shows the schematic diagram of such a generator. The variable resistor which is in series with the field winding is called a field rheostat and serves the purpose of adjusting the field current to a desired value. When the field current is taken directly from the generator output itself it is called self-excited. There are several possible connections for this purpose. Field and armature circuits may be in series, or in parallel, or there may be combinations of series and parallel connections.

4.3 D.C. GENERATOR

D-C generators produce a voltage in the manner discussed in Section 4.2; that is, the produce a sinusoidal voltage. However, a mechanical device called a commutator is used to provide a unidirectional current to the output terminals of the

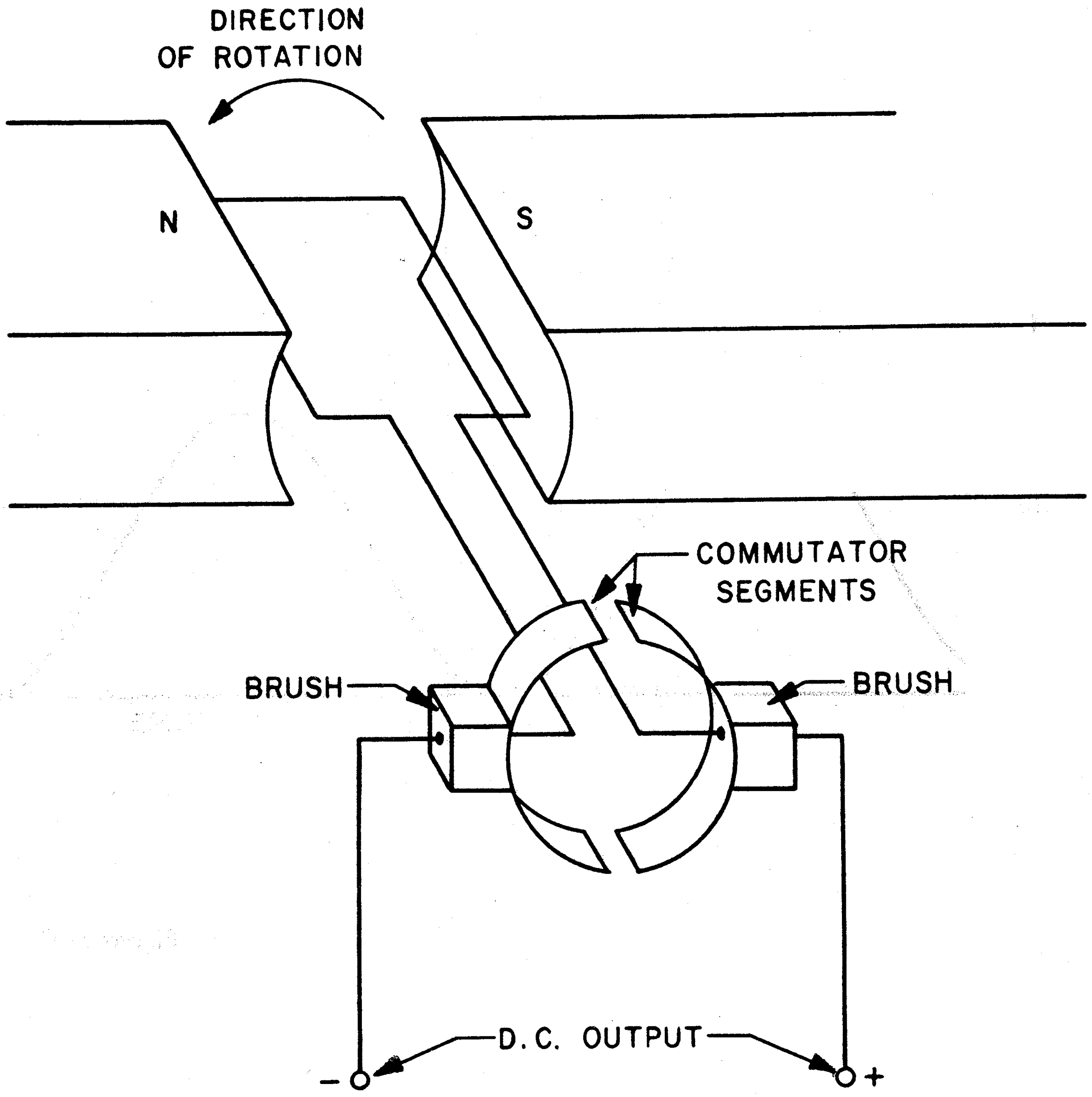

DIRECTION OF ROTATION
N
S
COMMUTATOR SEGMENTS
BRUSH
BRUSH
D.C. OUTPUT
−
+

Figure 3-62

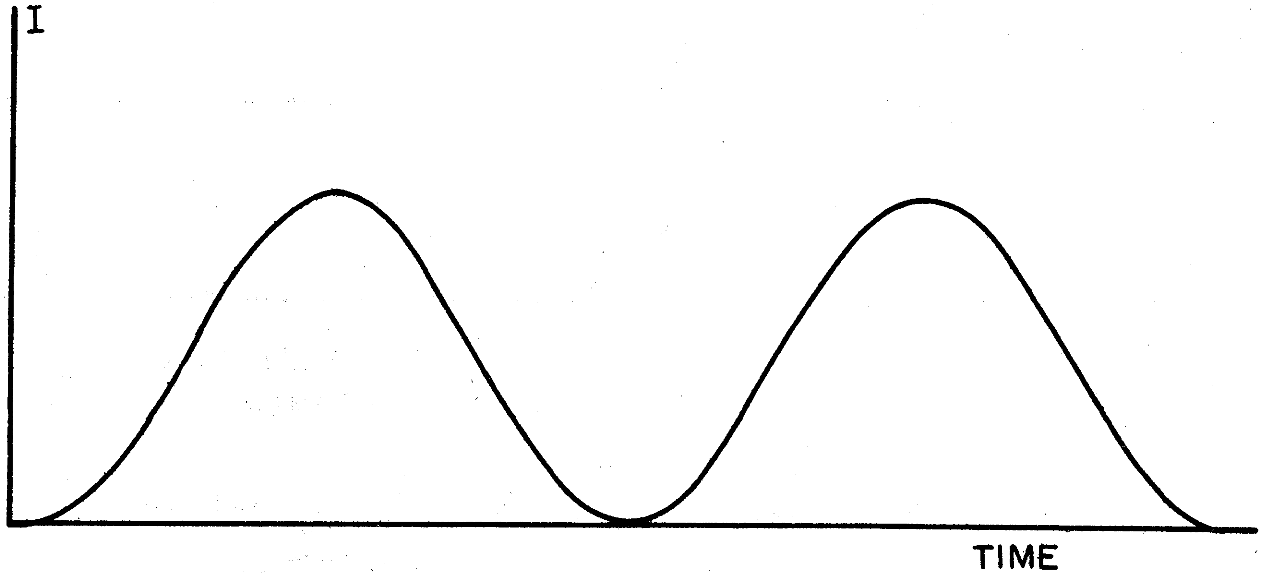


Figure 3-63

generator. In terms of the simple single loop generator of Section 3.2, the action of the commutator is as follows: Each end of the loop of wire is connected to a segment of the commutator. (See Figure 4-62.) The brushes or sliding contacts are placed so that electron flow from the loop is always directed into one brush while electron flow into the loop is always directed away from the opposite brush regardless of the direction of the electron flow in the loop. Therefore, the brushes carry electrons flowing in one direction only so that the output terminals of the generator are always of the same polarity. The load current has the same wave shape as the voltage between the brushes; it is unidirectional current. (See Figure 3-63.)

In a practical generator, the single loop armature shown in Figure 3-62 is replaced by many turns of wire which are wound on a cylindrical drum, or core. There are several methods of winding the wire to form an armature; however all these methods have in common the fact that the two ends of the armature wire are connected together so that the armature winding is a closed loop. As the armature is rotated in a magnetic field, the instantaneous emf of any of the individual turns depends upon its instantaneous motion with respect to the field. Thus, at any instant there is a turn which is moving parallel to the field from north to south and another turn which is moving parallel to the field from south to north

A line which passes through these two turns is defined as the neutral axis of the generator. All the turns of the winding above the neutral axis are cutting across the magnetic field in one direction, while all the turns of the winding below the neutral axis are cutting across the field in the opposite direction. Thus, the emfs induced in the turns above the neutral axis are in one direction, while the emfs induced in the turns below the neutral axis are in the opposite direction. Thus, the emfs in the turns above the neutral winding add, while those in the turns below the neutral axis add. On the other hand, when the total emf of the turns above the neutral axis is considered with respect to the total emf of the turns below the neutral axis, it is found that one emf tends to force current through the closed armature loop in one direction while the other emf tends to force current through the closed loop in the opposite direction. Since the two emfs are equal, the armature winding is in a state of equilibrium so that no current flows in it unless an external load is connected to it. This is analagous to two batteries of equal voltage having their positive terminals connected together and their negative terminals connected together. In such a case the batteries are said to be connected in parallel, because if a load is connected between their joined negative terminals and their joined positive terminals, half the load current will be furnished by one battery and the other half by the other battery. If the generator armature winding is to provide a d-c voltage

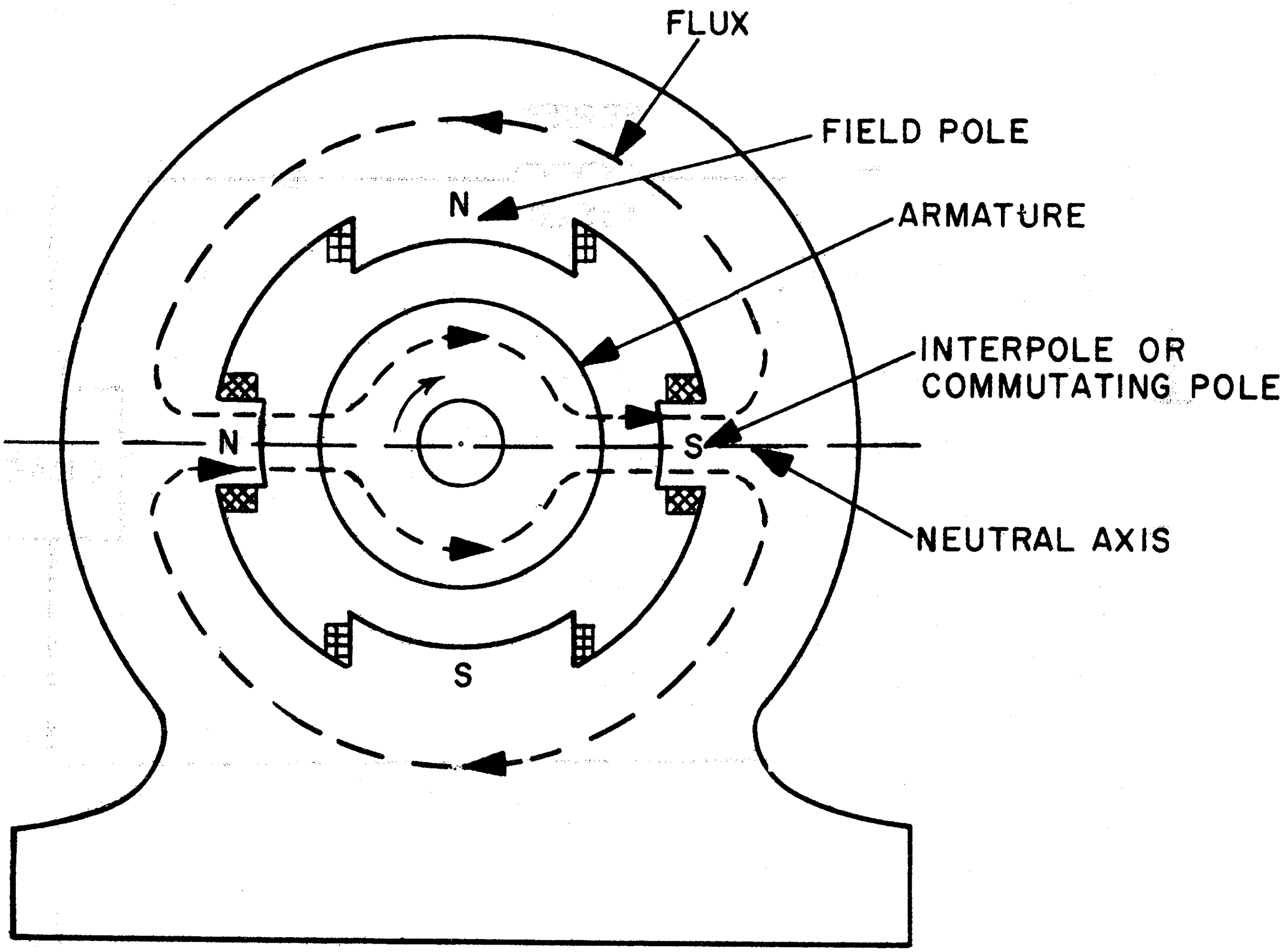

FLUX
FIELD POLE
N
ARMATURE
INTERPOLE OR COMMUTATING POLE
N
S
NEUTRAL AXIS
S

Figure 3-64

to an external load, an arrangement must be provided whereby one side of the load is always connected to the turn of the armature which is passing downward through the neutral axis and the other side of the load is always connected to the turn of the armature which is passing upward through the neutral axis (this being equivalent to connecting the load across two batteries connected in parallel). Such a connection can be realized if each of the turns has an associated slip ring segment, and two brushes are positioned so as to contact these segments as they pass through the neutral axis of the generator.

The practical generator which has been discussed up to this point is assumed to have only a single pair of poles. However, many generators are constructed with a number of pairs of poles. A multi-polar generator operates in much the same way as the single-pair generator just described; however, in the multi-polar machine there are as many neutral axes as there are pairs of poles, and thus a pair of brushes is provided for each pair of poles.

In order to prevent sparking at the brushes of a d-c generator, commutating poles or interpoles (small auxiliary poles) are placed between the regular field poles of the generator as shown in Figure 3-64. The sparking is due to what is called armature reaction. Armature current, which varies with the load, sets up a magnetic field in the armature core that reacts upon the main field of the generator set up by the field poles. Thus, a distortion of the main field is caused. This distortion or armature reaction delays the collapse of the field of the armature winding which results in the field not being completely collapsed when the brush and the segment are parted. Then, because the brush disconnects almost instantaneously there is a very rapid change of current in the armature coil, inducing a large voltage which causes sparking. In order to prevent this sparking the brushes should make contact with a pair of commutator segments just as the current in the coil connected to the segments is reversing direction (i.e. passing through zero) and should break contact with the pair of segments half a cycle later when the current in the coil is again reversing direction. The coils should therefore be in a neutral position where no voltage is being generated in them, which is normally midway between each pair of field poles. However, due to the distortion of the field the neutral position is advanced, which requires shifting the brushes forward a little with each increase in load. By placing the commutating poles between the field poles, armature reaction is overcome, thus preventing distortion of the field and the need for shifting the brushes.

One question that naturally arises in connection with a self-excited generator is as follows: Since the field current is supplied by the induced voltage and since no voltage can be induced unless a magnetic field is available, how does it get

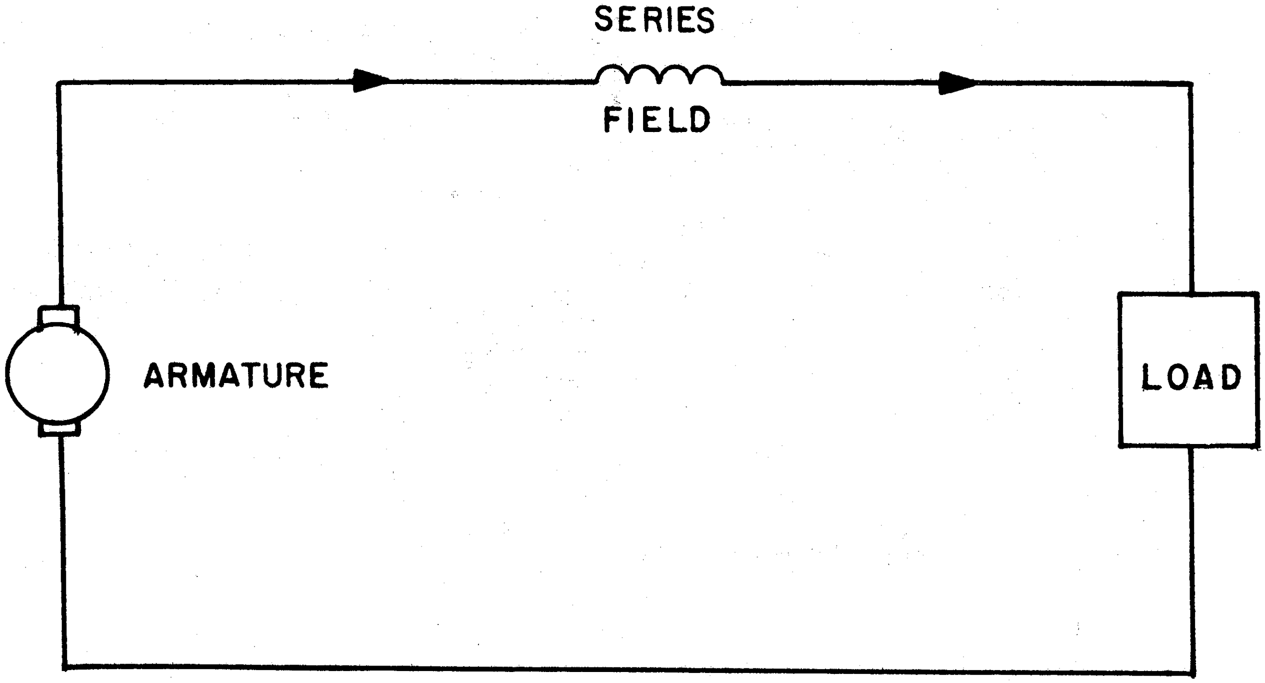


Figure 3-65

started? The answer is that a sufficient field for starting is ordinarily provided by the residual flux of the field poles. However, any of the following conditions may prevent the generator from building up to its normal voltage output:

a. Reversal of field connections
b. Excessively high field resistance
c. Excessively low residual field flux
d. Excessively slow armature speed
e. Brushes too far off neutral.

There are several different field arrangements for self-excited d-c generators. The particular field arrangement makes the generator either a series-connected generator, shunt-connected generator or a compound-connected generator.

4.3.1 Series-Wound Generator

A series-wound generator is one in which the field coil is wound with a few turns of heavy wire through which the entire current of the generator flows (see figure 3-65). The current rating of the series field should equal that of the armature. However, the resistance of the series field should be so small that its voltage drop is only a small per cent of the voltage induced in the armature. The voltage of the generator increases with the load because of the increase of current through the field winding. Series-wound generators are rarely used today because the voltage variation in response to even a small load variation is pronounced enough to be objectionable.

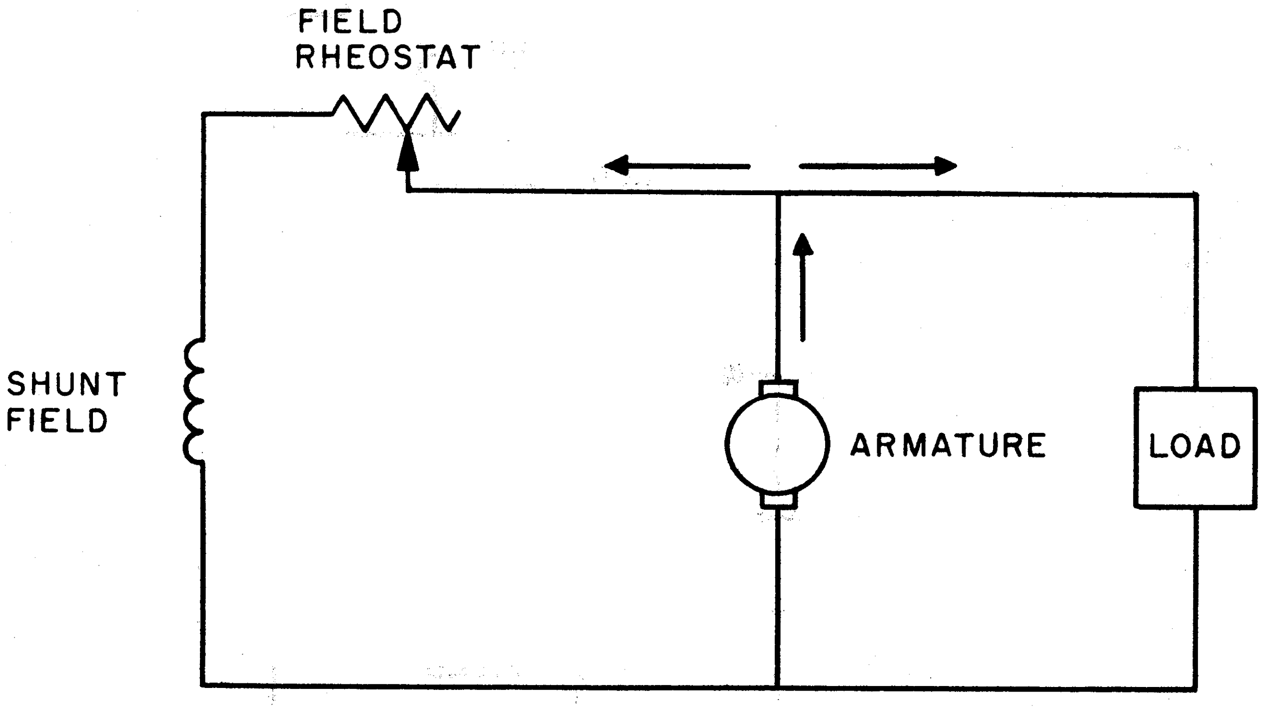
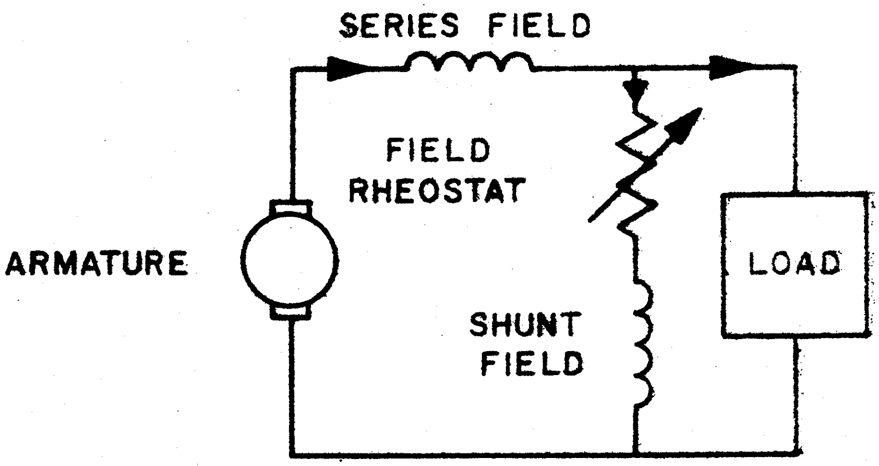
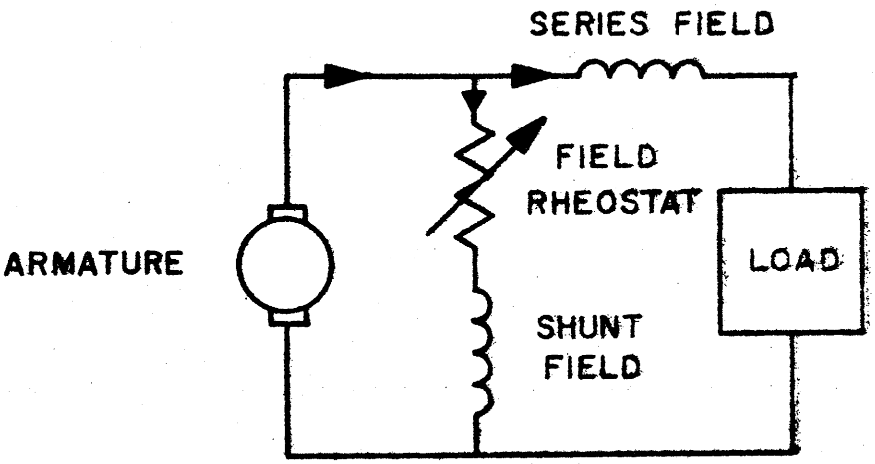


Figure 3-66



LONG SHUNT



SHORT SHUNT

Figure 3-67

4.3.2 Shunt-Wound Generator

A shunt-wound generator is one in which the field winding is connected across (in parallel with) the armature, as shown in Figure 3-66. In many instances the shunt field is designed to have a resistance somewhat below the exact value required, and is connected in series with a variable field rheostat so that the field circuit resistance and current may be adjusted to the desired value, which in turn adjusts the output voltage. The field of a shunt generator requires a current which must be only a fraction of the total current supplied by the armature. Shunt generators are used principally where an approximately constant load is to be maintained on the machine or where the circuit will, as the load increases permit the voltage to drop off slightly. In many cases they are also used for the charging of storage batteries and for the excitation of alternators.

4.3.3 Compound-Wound Generator

A generator with both shunt and series field excitation is called a compound generator. There are two possible connections, long shunt and short shunt as shown in Figure 3-67, A and B, respectively. The difference between these two types is small and unimportant. A compound-wound generator may essentially be considered as a shunt generator with the series field added to modify the external characteristics. Usually the voltage is controlled by means of a shunt field rheostat just as the shunt wound generator is controlled. The decrease of voltage with an increase in load, which occurs when only shunt fields are used, is compensated for by the increase in voltage due to the series winding. The voltage therefore remains practically constant when the generator is running at constant speed, regardless of the variation in load. When the series field winding is connected so that it strengthens the flux produced by the shunt field winding, it is known as cumulative compounding. It counteracts the shunt generator's falling off of terminal voltage under load.

The amount of compounding may be varied by changing the ampere turns of the series field for a given generator current. This is accomplished by changing the number of turns in the series field or by putting a variable resistor in parallel with the series field. If the number of series turns is chosen correctly, the full-load voltage is equal to the no-load voltage. Such a compound generator is called flat compounded.

A cumulatively compounded generator with a number of series turns higher than that required for flat compounding is called over-compounded; one with fewer series turns is under-compounded. In an over-compounded generator, the increased IR drop through the external circuit leads, caused by an increase in load current is compensated for by an increase in generator output voltage. This provides a constant voltage at the load center, which is a highly desirable condition. Compound-wound generators are used in most direct-current light and power installations.

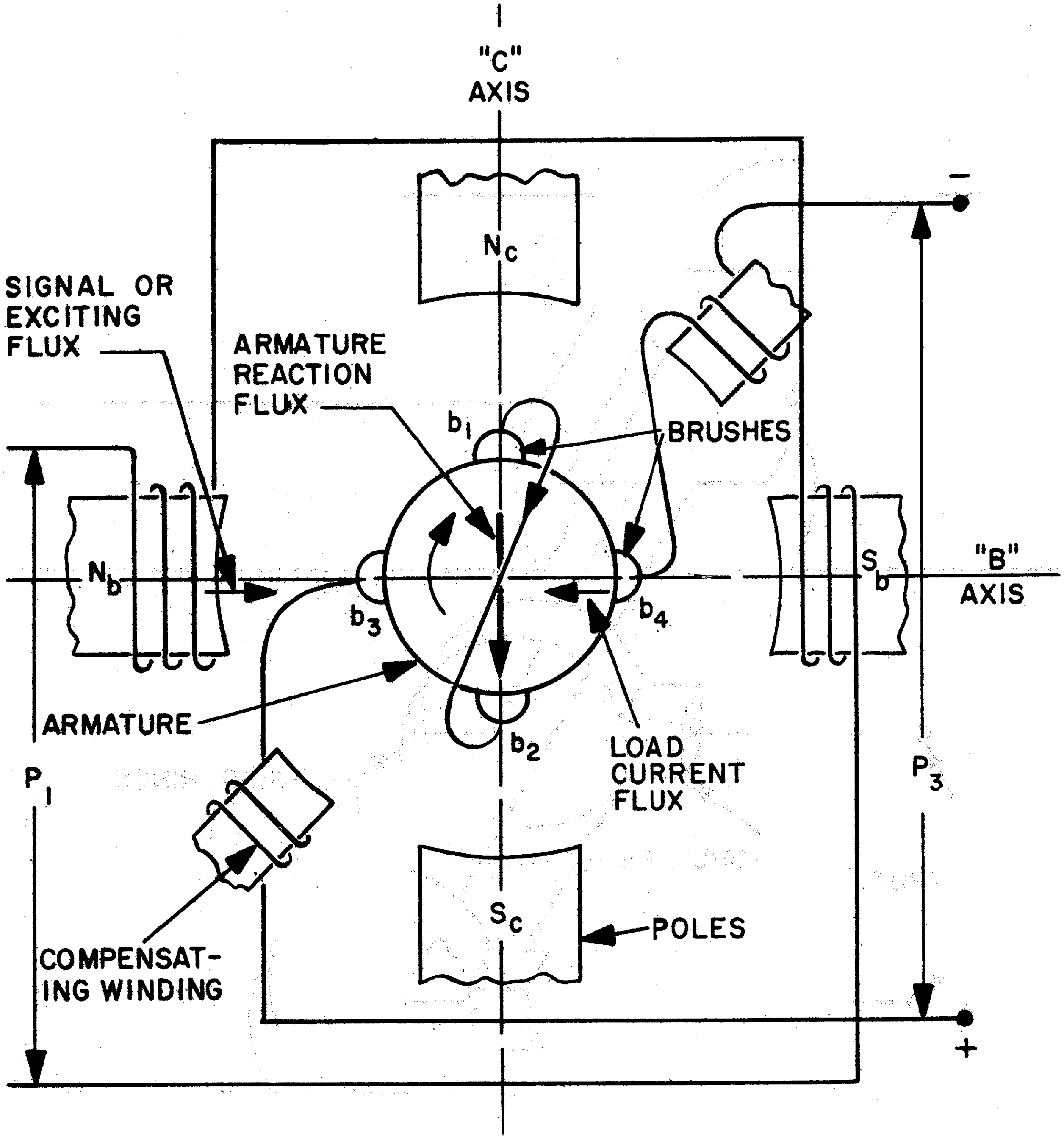

"C" AXIS
Nc
SIGNAL OR EXCITING FLUX
ARMATURE REACTION FLUX
b1
BRUSHES
Nb
b3
b4
Sb
"B" AXIS
ARMATURE
b2
LOAD CURRENT FLUX
P1
P3
Sc
POLES
COMPENSATING WINDING
−
+

Figure 3-68

4.3.4 Amplidyne Generator

The amplidyne generator is a device which combines, in one machine, the equivalent of two stages of amplification. The signal power P, is used to excite the field winding of poles N_B and S_B (see Figure 3-68). Excitation of these poles produces flux through the "B" axis of the generator. A voltage is induced by this flux between the brushes b_1 and b_2. If these brushes are short-circuited, a large current flows in the armature resulting in a large flux through the "C" axis by armature reaction. In order to reduce the reluctance of the flux path along the "C" axis a second set of poles is provided on the "C" axis. The "C" axis flux generates a voltage between the main brushes b_3 and b_4 which serve as the output terminals.

This generator is in effect two machines, one on the "B" axis and the other on the "C" axis, but both use a common armature which is wound for two pole operation. The stages of amplification are obtained in this generator with an amplification factor of 10^4 or better.

A compensating winding, which is stationary and supported by the poles is used to neutralize the armature reaction effects of the load current opposing the control flux on the "B" axis. The load current of the armature is carried by this compensating winding. It must have the same number of ampere turns as are produced in the armature by the load current. The

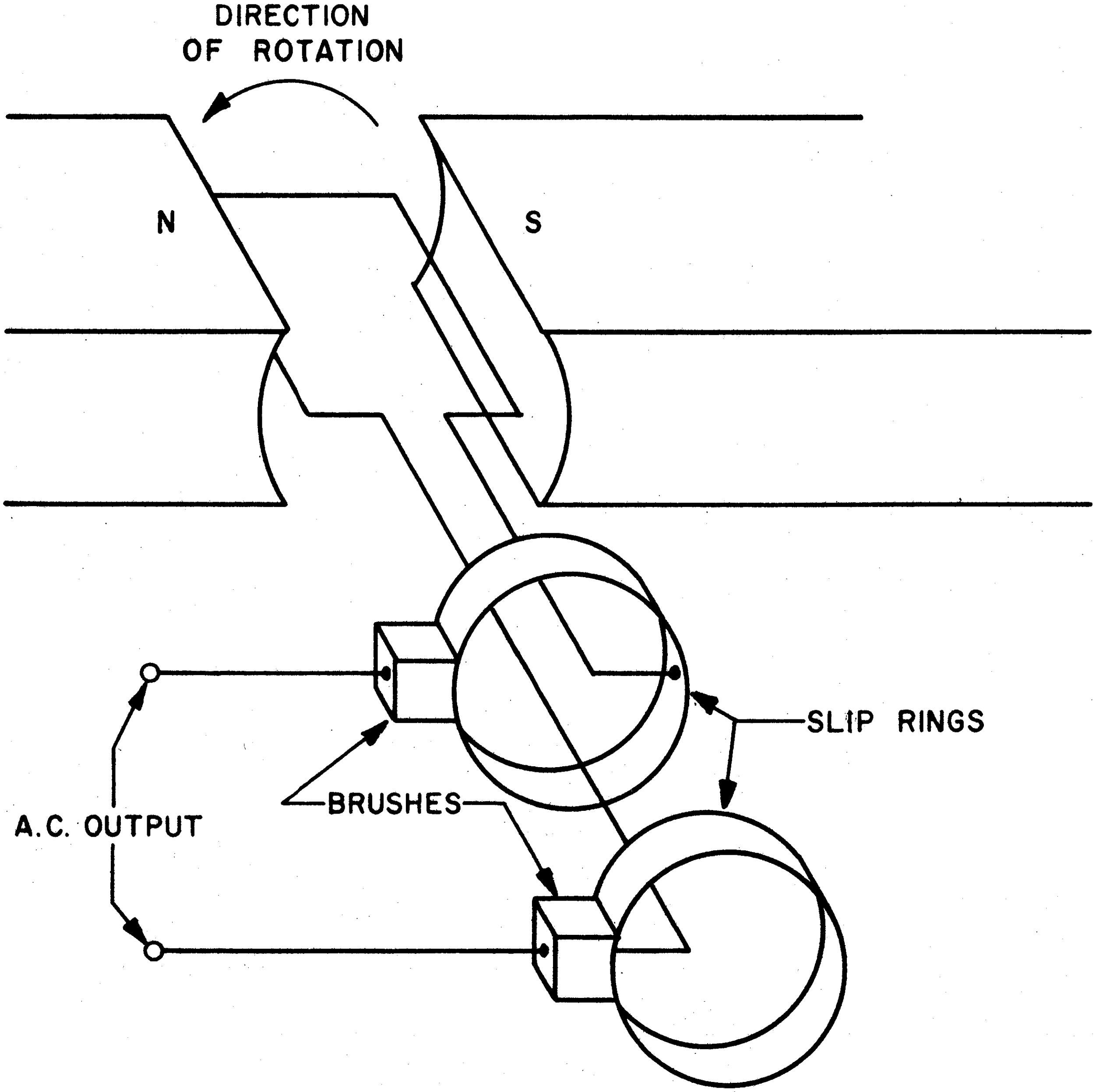


Figure 3-69

direction of the magnetizing action of the compensating winding is made opposite to the demagnetizing action of the armature reaction so that the two effects cancel each other out. The complete output circuit current compensation by the compensating windings is what particularly distinguishes amplidynes from other cross-field machines. This compensation allows the output voltage to be directly proportional to the exciting voltage.

Some of the advantages of an Amplidyne generator over a tube amplifier are: high output power (up to 100 kw or more); output is at low voltage suitable for operating motors; control by a low excitation power (a few watts); efficiency in transferring power from motor driving amplidyne to electrical load may be quite good (70 per cent or more).

4.4 A-C GENERATORS

A-C generators produce a voltage in the same manner as d-c generators, as explained in section 4.3. However, the ends of the armature coil are terminated at slip rings, which make continuous contact with fixed brushes as they rotate (see Figure 3-69). Therefore the polarities of the brushes and terminals alternate between positive and negative, and any load connected to the terminals draws alternating current.

There are two general types of a-c generators (also called synchronous generators or alternators) in use today. They are, the revolving armature type and the revolving field type. The revolving armature type of generator is very similar to a d-c generator, having stationary magnetic fields and the revolving armature. The revolving armature type of a-c generator is usually confined to low voltages and very small capacities. This is due to the difficulty of insulating the slip rings and brushes for high voltage and the large number of brushes required for heavy currents. The revolving field type of a-c generator is the one used mostly today. It consists of a stationary armature and a field which rotates. The rotating and stationary members of a revolving field type of generator are frequently referred to as the rotor and stator respectively. The slip rings and brushes are only required to supply low voltage exciting current for the fields and are easily insulated. The revolving fields are divided into two classes; the first type consists of definite polar projections around which the field coils are wound. This is called salient pole construction. The second type is distributed field winding construction, where the field coils are imbedded in the face of the revolving member in a distributed winding and there are no polar projections.

Generation of power at commercial frequencies - 25, 50, 60 and 90 cps (cycles per second) is accomplished almost exclusively by the use of synchronous machines. Although this type of machine is also used as a motor the discussion here is confined to the use of the synchronous machine as a generator. The characteristics for motor operation are discussed in section 4.8.2 .

A small d-c generator, called an exciter is often used to furnish the field current in a synchronous generator. Since the field must be excited by a direct current a synchronous generator cannot be self-excited by connecting its field directly to the alternating terminal voltage.

4.5 MECHANICAL ENERGY

When a conductor carries an electric current while in a magnetic field a mechanical force is exerted on the conductor. The direction of this force can be determined by reference to Figure 3-21. The electron flow within the conductor is shown by a cross, directed away from the reader and hence using the left hand rule as explained in Part 3, Chapter 2 on magnetism, the magnetic flux which it produces is counterclockwise. The lines below the conductor produced by the poles and by the conductor are both to the right, therefore strengthening the resultant field, but above the conductor the two fields oppose each other creating a weak resultant field. The wire is pushed upward as a result of this flux distribution. This happens because the magnetic lines of force contract to their shortest length (tension in lines) and push sidewise upon each other (lateral repulsion of lines). A more mechanical rule which is used to determine the direction of the force is the "Right-Hand Motor Rule". This rule, states: to determine the direction of force extend the thumb and first two fingers of the right hand at right angles to one another. If the forefinger points in the direction of the magnetic lines of force and the middle finger in the direction of electron current in the conductor, the thumb points in the direction of the force of the conductor.

4.6 ELECTRIC MOTOR

Since motor action is the reverse of generator action, any device that works as a generator, can also operate as a motor. An electric motor is a machine for converting electrical energy into mechanical energy. Therefore a motor is a machine which is connected to an electric voltage source instead of being driven by a prime mover. This reversibility applies to direct or alternating machines. In some cases, the machines may not work equally well as motors or generators, and modification of design for motor operation may be desirable. However, the difference between motor and generator design in d-c machines is not important.

4.7 D-C MOTOR

Figure 3-62 serves to illustrate the operation of a direct-current motor as well as a generator. If the outside circuit has a voltage that causes a current to flow in the armature loop a magnetic field is produced in the conductors. There is also a second magnetic field produced by the N and S poles. The north pole of the main magnetic field attracts the south pole of the armature, and, since the loop is free it revolves. However, the instant that the north and south poles come opposite, the commutator reverses the current in the armature, making like poles opposite and the loop is then repelled and

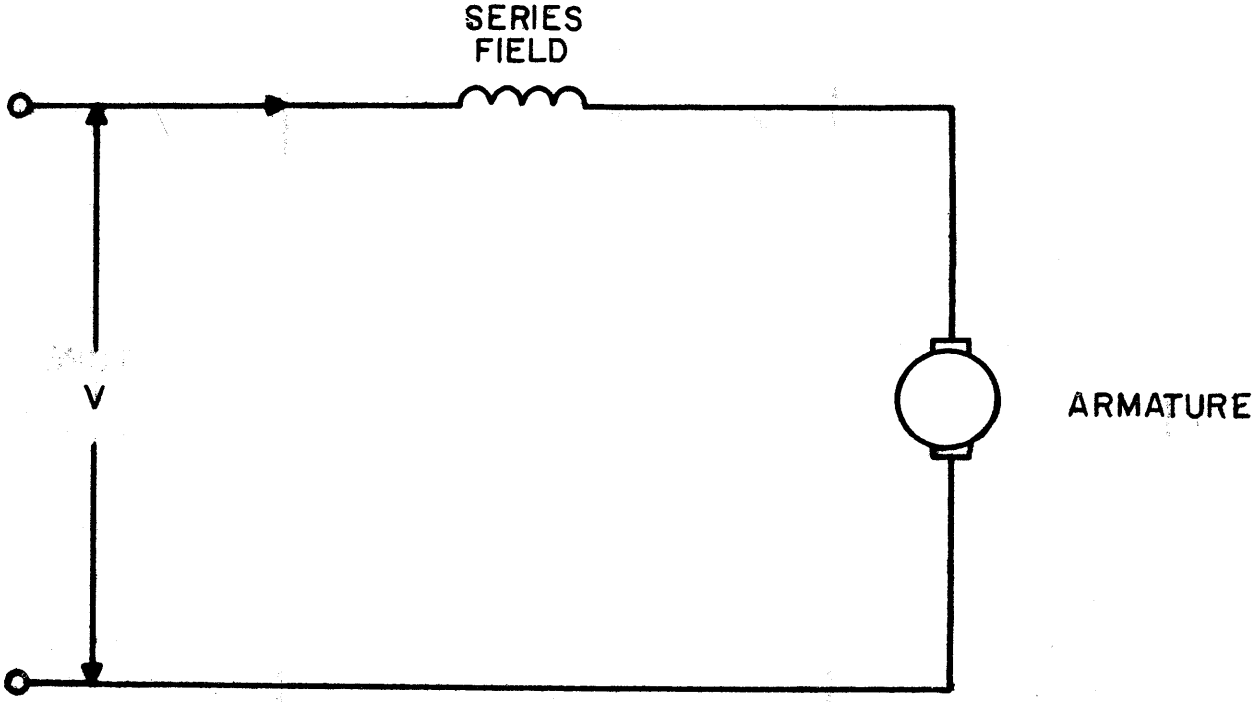


Figure 3-70

forced to revolve further. Each time unlike poles approach, the armature current is reversed. This continues as long as there is current in the armature and field windings and the armature is free to revolve. In actual motor armatures there are many loops, called armature coils, each with its terminals connected to adjacent commutator segments. Therefore the attracting and repelling action is more powerful and more uniform instead of the weak and jerky action obtainable with the single loop armature.

There are three general types of d-c motors, as there are three general types of d-c generators. They are the series, shunt and compound, which are named according to the way the field is connected.

4.7.1 Series-Wound Motor

A series-wound motor, Figure 3-70, is one in which the field coils and armature are connected in series and the entire current flows through the field coils. Since the armature current flows through the field coils they require few turns. The turns must be made of heavy wire so that field coil resistance is as small or smaller than that of the armature.

The speed of a series motor varies with the load and is usually governed by varying the amount of resistance in series with the armature and field. With a light load it runs faster than with a heavy load, and with each increase in load it slows down. However, it carries a very heavy overload before

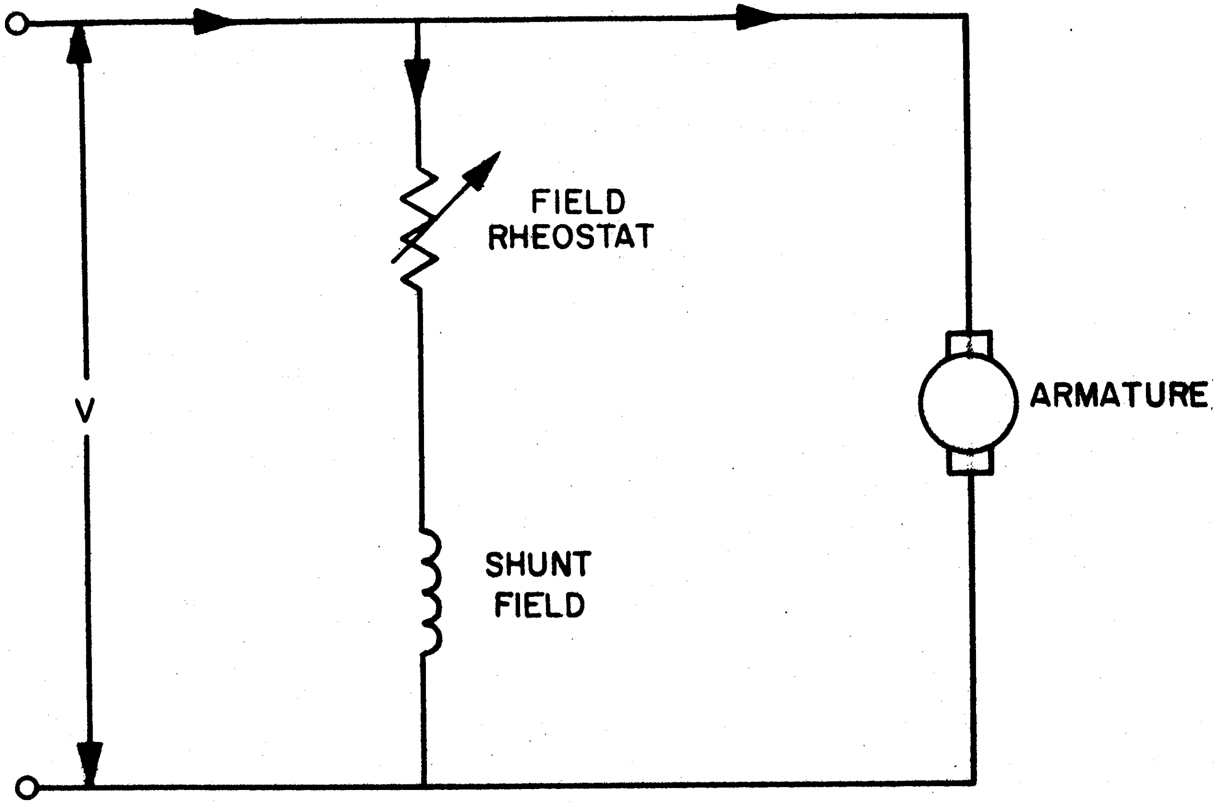


Figure 3-71

being stalled. Due to the ability of series motors to start a heavy load with less current than other motor types they are almost universally used for heavy service requiring frequent starting.

Series motors are never belted to the driven machines, because if the belt breaks it removes the load entirely causing it to "race" at a dangerously high speed. They are almost invariably directly coupled or geared to the load. Series motors are used generally where intermittent operation is required and where starting under heavy loads is encountered. They are used in applications where their operation is continually monitored by an operator.

4.7.2 Shunt-Wound Motor

A shunt-wound motor, Figure 3-71, is one in which the field coils and armature are connected in parallel, and in which the field coils take only a small portion of the line current.

Regardless of the load a shunt motor runs at nearly the same speed, even though greatly overloaded. These motors are generally classified among the constant-speed motors. Shunt motors do not run away under no-load conditions so they may be belted to a line shaft which drives several machines, or it may be used to drive only a single machine, being connected to it directly by a coupling, or through a belt, chain or gear. Their maximum torque is considerably less than that of a series motor. Therefore, they are used where constant speed

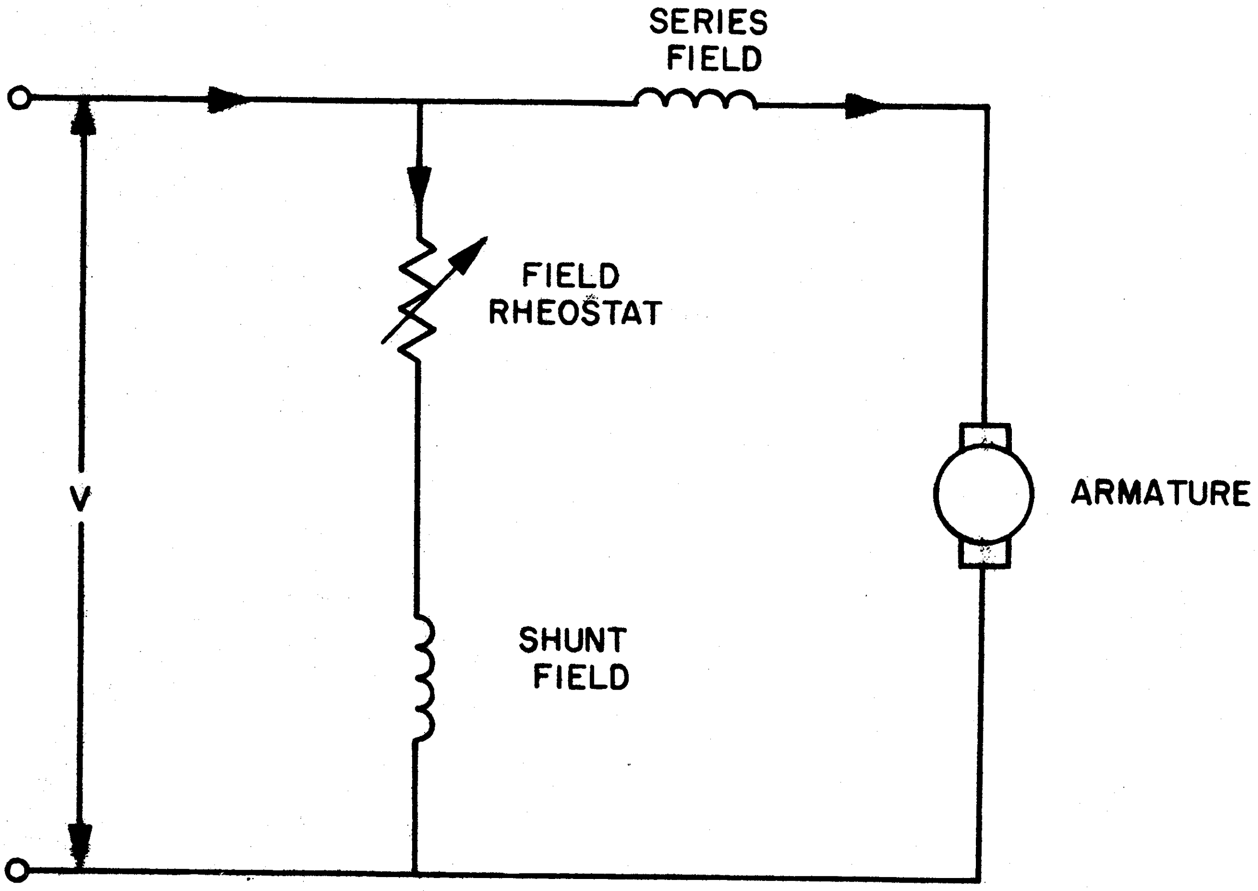


Figure 3-72

and low starting torque is necessary.

4.7.3 Compound-Wound Motor

A compound-wound motor has both series and shunt field windings as shown in Figure 3-72, giving it a combination of the characteristics of both series and shunt machines. As the load changes, the speed changes, but it does not change as much as in a series motor and it usually changes a great deal more than a shunt motor. A maximum speed, faster than which the motor does not run even though the load is thrown off entirely, is fixed by the shunt winding. The series winding causes the motor to drop in speed and to pull harder as the speed drops. Similar to the series motor, it has excellent characteristics for starting heavy loads and still it is in no danger of "running away" at light loads. Compound wound motors combine large starting torques with fairly constant speed.

Compound motors should always be used wherever the line voltage fluctuates suddenly and widely. They are usually designed so that there is a drop of about 20 per cent in the speed between no load and full load. In order to meet special conditions they may be wound for greater or less speed change.

4.8 A-C MOTORS

The most commonly used alternating current motors are divided into two general classes according to their principle of operation; namely, induction and synchronous. Induction motors are further divided into two general classes; squirrel cage and slip ring or wound rotor motors. A-C motors are also classified as single and polyphase motors. These general types have many modifications, both electrically and mechanically, which make it possible to secure motors for almost any given application.

4.8.1 Induction Motor

The induction motor corresponds quite closely to a transformer, because the magnetic flux produced provides transformer action between a stator and a rotor. The stator usually serves as the primary and the rotor as the secondary, although at times it may also function as the primary. Therefore, an induction motor may be considered as a transformer with a movable secondary. The motor action is due to the dragging effect. This dragging effect is the result of the rotating flux (produced by the current through the windings of the stator) inducing a current in the rotor (conducting body) which in turn produces the magnetic field that opposes the original rotating flux. This transformer action between stator and rotor removes the need for electrical connections to the rotor. Because of this important characteristic the need for moving contacts either in the form of slip rings or commutators is eliminated. This avoids an important source of trouble and maintenance and gives the induction motor its outstanding reliability.

The induction motor is also one of the most simple in construction and has the broadest application making it by far the most commonly used of any of the types mentioned. A serious limitation of the induction motor is that it is essentially a constant-speed device and does not offer the speed adjustment provided by field control of a shunt motor. Large induction motors are always polyphase machines and are usually of the 3-phase type. The single-phase types are usually of the fractional horsepower sizes. The squirrel cage motor which derives its name from the fact that the rotor winding resembles the wheel of a squirrel cage, is the most common form of the induction motor.

4.8.1.1 Squirrel Cage Motor

The squirrel cage induction motor is a general purpose motor. The first requisite is that its mechanical construction be sturdy and rigid to enable it to operate satisfactorily with minimum maintenance under the severe service conditions which are often encountered.

The squirrel cage induction motor is one of the most efficient motors built, being exceeded in this respect only by the synchronous motors in large sizes. Under normal load and voltage conditions the speed of the squirrel cage motor is nearly constant. The difference in speed, between the synchronous speed and the load speed for any given load is called the slip. The synchronous speed is the speed of the rotating flux. The speed of the rotating flux per minute is directly proportional to the frequency and indirectly proportional to the pairs of poles and can be determined by using the following formula:

$$\text{Rotating Flux/minute} = \frac{\text{60 x FREQUENCY}}{\text{Pairs of Poles}}$$

Slip is also a measure of the losses in the rotor windings. There is a definite relationship between slip and the efficiency of the motor. The higher the slip the lower the efficiency.

The efficiency and power factor of an induction motor is lower at light loads than at full loads. In selecting a motor for a definite load, the size should be such as to operate at nearly full load. High speed motors have higher power factors than low speed motors. Therefore, in making the selection of a motor the size and speed should be carefully considered, to give the most satisfactory and economical arrangement with the driven machine.

The stator of a squirrel cage induction motor, or for that matter, any induction motor, contains conductors which are placed in slots on the stator. These slots serve to shorten the air gap and to protect the insulation of the winding from damage. If two additional windings are added to the unused portions of the stator of a single phase machine the output of the machine is trebled with relatively little additional material and labor. Each of the three windings are actually a separate machine, but in its entirety it is called a three phase induction motor. Although the current in the three windings may be equal they are

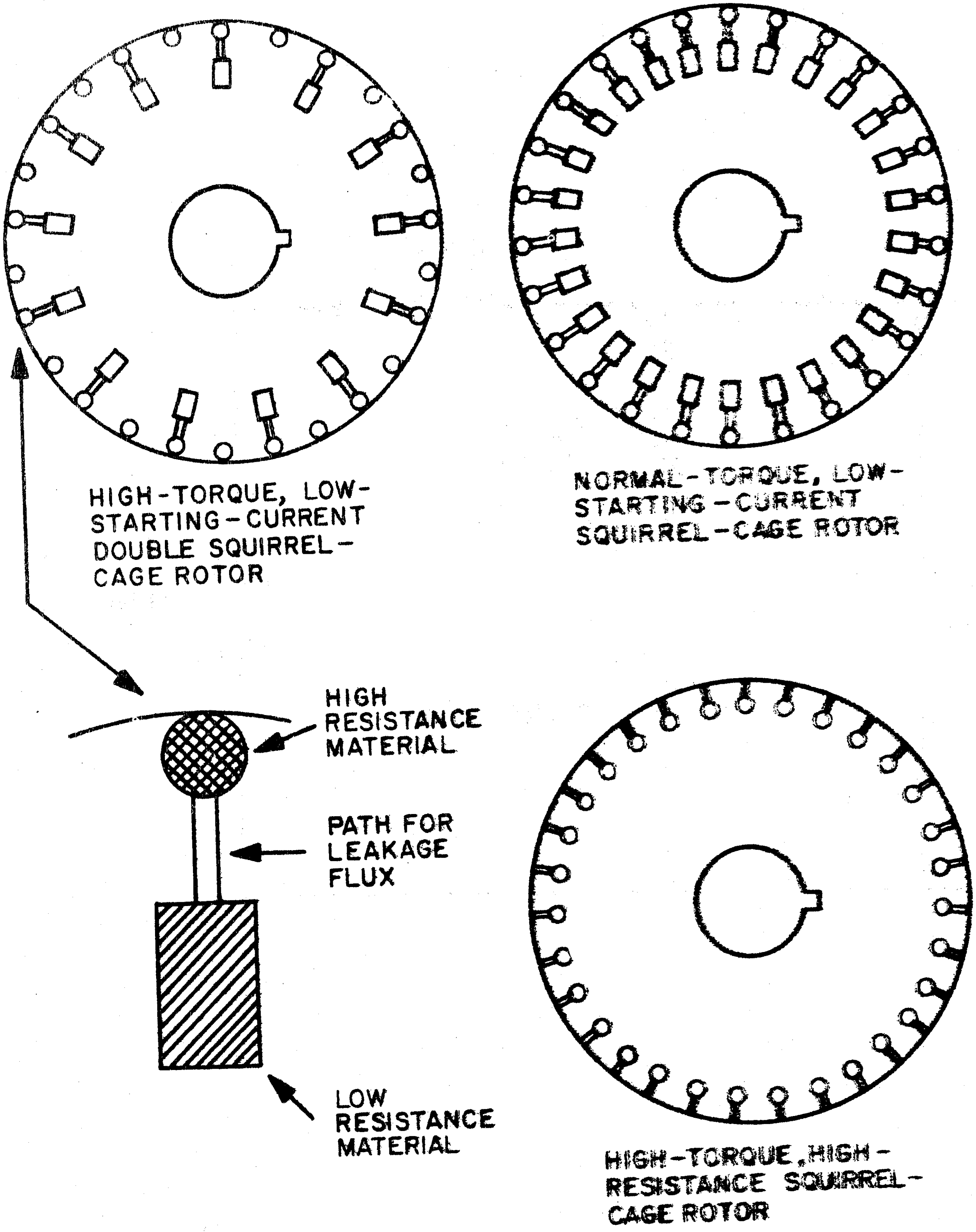

HIGH-TORQUE, LOW-STARTING-CURRENT DOUBLE SQUIRREL-CAGE ROTOR
NORMAL-TORQUE, LOW-STARTING-CURRENT SQUIRREL-CAGE ROTOR
HIGH RESISTANCE MATERIAL
PATH FOR LEAKAGE FLUX
LOW RESISTANCE MATERIAL
HIGH-TORQUE, HIGH-RESISTANCE SQUIRREL-CAGE ROTOR

Figure 3-73

separated by 120 degree phase angles, which is due to the location of the windings on the stator.

The squirrel cage rotor is one of the simplest and cheapest rotors used. The rotor is composed of a group of conducting bars placed in slots in the laminations and connected together by rings at each end of the cylindrical rotor which solidly short-circuits the conductors. Figure 3-73 shows some typical rotor laminations for induction motors. Squirrel cage is a name derived from the fact that these conductors are buried in the iron laminations and form a cage like appearance. The conduction bars and end rings are formed of an aluminum alloy which is cast in the laminations after they are stacked and fastened to the shaft. The rotor conductors are also arranged in either 1, 2, or 3-phase and may contain the same number of phases as the stator associated with the squirrel cage induction motor.

4.8.1.2 Slip Ring or Wound Rotor Induction Motor

A slip ring or wound rotor induction motor differs from the squirrel cage motor only in that power applied to the rotor windings causes the fields to be induced in the stator. The terminal connections of the rotor windings are brought out to slip rings and leads from the brushes on the slip rings are brought to an external resistance which can be regulated at will. The principle of operation is the same as that of the squirrel cage motor. This motor operates at high efficiency and with a low slip.

4.8.2 Synchronous Motor

A synchronous motor is one which rotates at the same speed as the alternator that supplies its current, or at a fixed relationship to that speed. It does not vary greatly from an alternator in construction, having the stationary armature and revolving field. Generally the field coils are excited by direct current from a small direct current generator.

The stator produces a rotating field as previously explained for induction motors. The rotor field has a d-c voltage applied to it so that it acts like a permanent magnet on a pivot. This d-c energized magnet follows the rotating field which produces the torque of the motor.

The synchronous motor is not self-starting. The torque reverses with the same frequency as the current and the inertia does not give the motor a chance to start. In order to make the motor self-starting a squirrel cage (amortisseur) winding is added which is effective both in starting and in preventing hunting speed characteristics (motor oscillates in phase position and therefore fluctuates in speed above and below its synchronous speed). When no amortisseur winding is furnished on the motor a starting motor must be furnished which is powerful enough to get the machine up to synchronous speed where the torque of the synchronous motor becomes continuous and effective.

Synchronous motors possess several unique characteristics. Any variation in load does not cause a corresponding change in speed, as with the induction motor, because this machine operates at a speed which is set by the frequency of the power supply. The frequency of the supply circuit and the number of poles determine the speed. The speed is found by dividing the frequency in cycles per minute by the number of pairs of poles, the quotient being the revolutions per minute.

In comparison to the corresponding induction motor of the same rating, the efficiency of the synchronous motor is somewhat higher. This is especially true with low speed motors.

4.9 CONTROLLERS

Controllers as the name implies are used to protect, start, stop, and regulate the speed of motors and generators. They are usually found at the immediate source of motor or generator or at some control panel. There are many varieties of controllers some of which are: starting boxes, speed controllers, circuit breakers, and contactors.

4.9.1 Starting Boxes

4.9.1.1 D-C Starting Box

A starter is necessary for all but very small direct-current motors. Any good motor must have low resistance, in order to use electricity for the generation of power rather than for heat. However, if a motor with low-resistance windings is connected directly to the service wires while the armature is

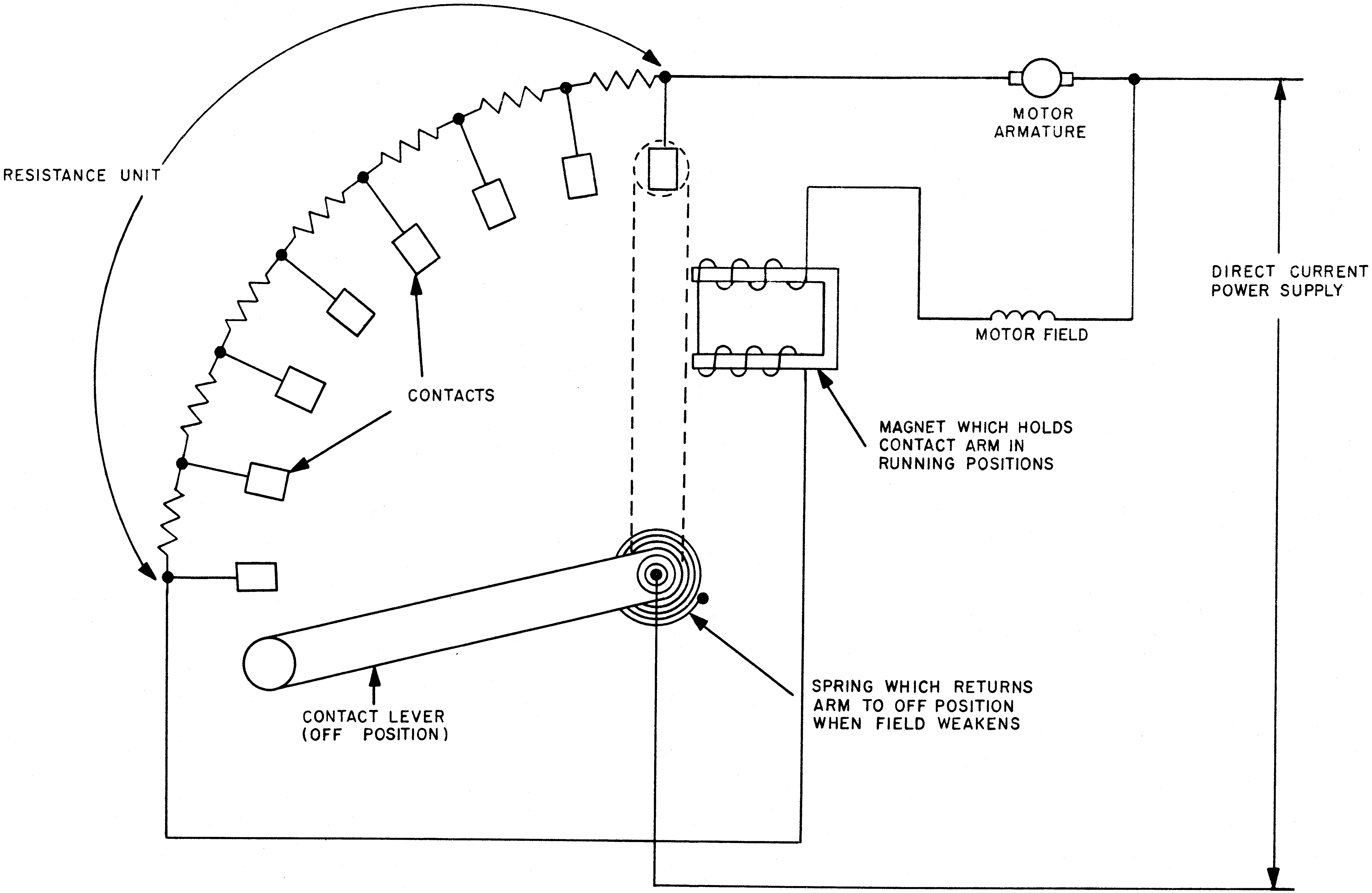

MOTOR
ARMATURE
RESISTANCE UNIT
DIRECT CURRENT
POWER SUPPLY
MOTOR FIELD
CONTACTS
MAGNET WHICH HOLDS
CONTACT ARM IN
RUNNING POSITIONS
SPRING WHICH RETURNS
ARM TO OFF POSITION
WHEN FIELD WEAKENS
CONTACT LEVER
(OFF POSITION)

Figure 3-74

standing still (not generating a back EMF), it takes a very large current. Therefore, in order to cut the current down to a reasonable amount a starter with considerable resistance is placed in the circuit. This resistance is cut out in sections (see Figure 3-74) as the motor gradually comes up to its normal running speed. A starter of this type is called a "starting rheostat" or "starting box". It not only protects the armature from excessive starting current, but prevents the motor from loading down the power source during starting. Current flowing from the switch arm passes through all of the resistances if the arm is in contact with the first plate. As the arm is moved from plate to plate the resistance gradually decreases until the last plate is reached, which means that the armature is connected directly across the line. The contact arm is actuated by a spring so that when the circuit is opened, it returns to the off position automatically.

4.9.1.2 A-C Starting Box

The most common method of reducing the starting current to squirrel cage induction motors is to reduce the applied voltage by means of a reduced voltage starter. The two general types used are the resistance and auto-transformer or compensator types.

The resistance type employs a large resistor in series with the line. A reduced voltage is produced at the motor terminals due to voltage drop across this resistor. Therefore the starting current of the motor is reduced in direct proportion to the reduction of the voltage applied to the motor terminals.

The compensator starter, in order to reduce the voltage to the motor terminals usually employs two auto-transformers. This type of starter has the advantage of drawing less current from the line for a given reduction in voltage tothe motor terminals.

In addition to manual starters there are also automatic starters. In an automatic starter the starting resistors are cut out by contactors called accelerating contactors, the operating coils of which are controlled according to the accelerating method used. The three common methods are: (1) counter-emf acceleration; (2) definite-time limit; (3) current-limit acceleration. In the counter-emf method, the operating coils of each accelerating contactor are connected across the armature terminals, and adjusted to close in succession as the counter-emf rises; in the definite-time limit method, the resistors are cut out in sequence, at the end of definite time intervals, regardless of the motor speed, current, or load; in the current-limit method the accelerating contactors operate in sequence, in direct response to the variation in armature current.

4.9.2 Speed Controllers

Speed controllers are constructed practically in the same manner as starters except that the resistance coils are made large enough to carry the current continually without over-heating. The low-voltage release attachment is made so that it holds the switch arm on any one of the contacts as long as normal voltage is on the line and releases it whenever the voltage fails.

Another method of speed control used mainly in connection with adjustable-speed shunt motors is by varying the magnetic field current. This is done by placing more resistance in series with the field coil, in the form of a rheostat. Speed is inversely proportional to the flux, therefore, adding resistance to the field winding, raises the speed. This method of speed control is efficient, because by the addition of the field rheostat resistance the I^2R losses in the field circuit are actually reduced. Except for very low speeds, a fairly wide range of speed variation can be obtained.

This method (field resistance) is also used for speed control of compound motors, through for any speed setting there is come variation in the speed with the load, whereas in the case of the shunt motor the speed is not significantly affected by load. For series motors the method is often used in modified form; by shunting part of the current around the field coils the magnetic field is weakened and the speed increased.

Speed regulation for induction motors is accomplished in three ways. They are: (1) inserting resistance in the rotor circuit; (2) changing the number of poles; (3) impressing voltage across the external terminals of the rotor slip rings.

Of the three methods mentioned the third method is the one most often utilized. When it is necessary to reduce the speed of the motor a voltage is impressed across the terminals of the rotor slip rings which is directly opposite in phase to the emf induced in the rotor. Thus, the current in the rotor is decreased, and in order to overcome the opposing torque, the rotor increases its slip (reduces it speed), causing the rotor current to increase to an amount sufficient to overcome the opposing torque. The greater the impressed counter-voltage the more the speed is reduced. In order to increase the speed of an induction motor a voltage, which is in phase with the induced emf in the rotor is impressed across the terminals of the rotor slip rings. When the impressed voltage is made exactly equal to the rotor emf the motor runs at its synchronous speed. By impressing a voltage which is greater than the necessary rotor emf, the motor operates at a speed greater than the synchronous speed.

A synchronous motor, on the other hand, always operates at the synchronous speed and does not experience any speed change with a change in load. Therefore it is usually not necessary to regulate the speed of a synchronous motor.

4.9.3 Circuit Breaker

Circuit breakers are devices which are used to protect various equipment from overload conditions and sometimes as equipment disconnect switches. This is accomplished by manually or automatically disconnecting the equipment from the

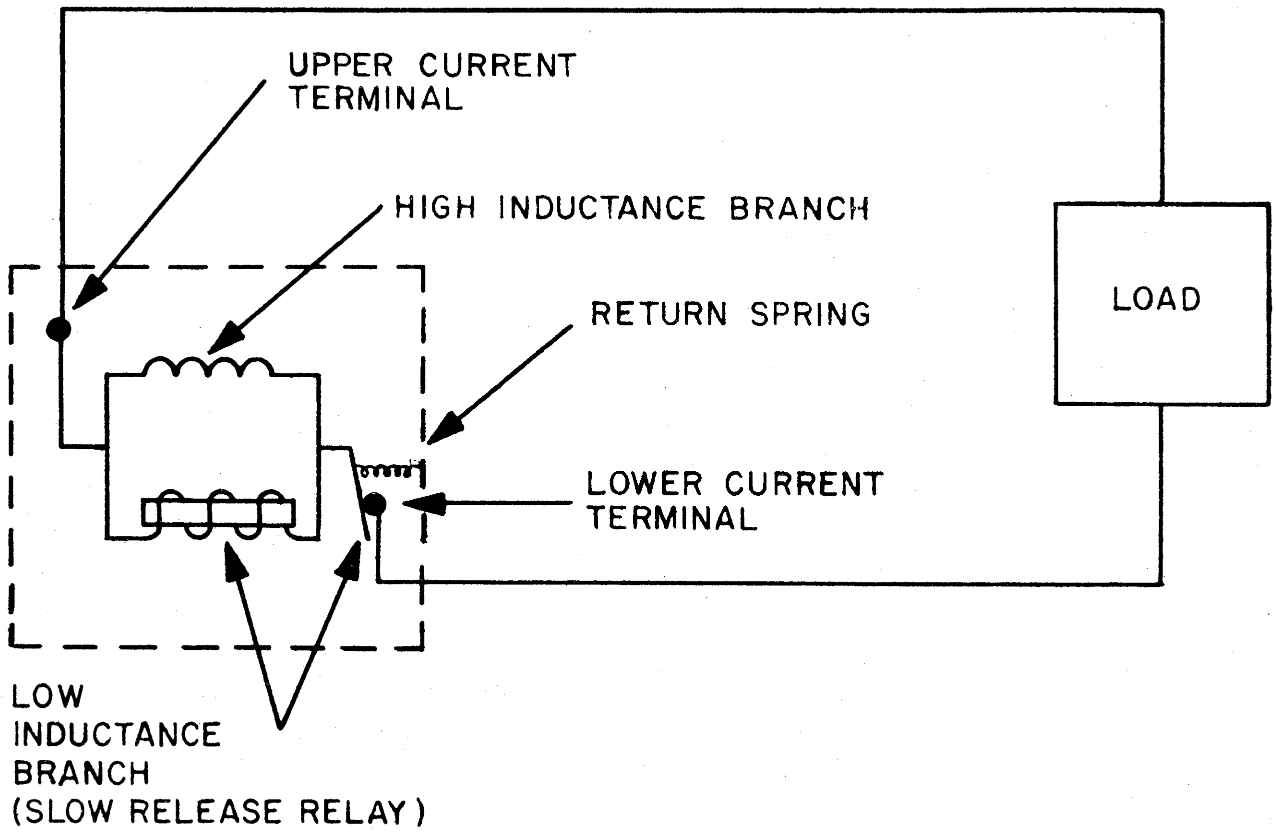

UPPER CURRENT TERMINAL
HIGH INDUCTANCE BRANCH
RETURN SPRING
LOAD
LOWER CURRENT TERMINAL
LOW INDUCTANCE BRANCH (SLOW RELEASE RELAY)

Figure 3-75

circuit when the current reaches a predetermined dangerous value or when shutdown is desired. Automatic reclosing circuit breakers are used so that interrupted circuits do not continue in this state unnecessarily. A mechanism operates to reclose the breaker after tripping, provided that the short circuit that existed has been repaired. However, if the short circuit still exists the breaker does not close.

In a typical circuit breaker the short circuit current may be anticipated by the breaker tripping mechanism. As shown in Figure 3-75 the bottom current lead of each pole is divided into two parallel branches, the lower or non-inductive branch (slow release relay) and the upper or hightly inductive branch. Since the two branches are in parallel the current divides inversely as the resistances, tripping the breaker according to the pre-arranged calibrations for the relay. When the current rises with great rapidity most of the current flows through the lower or tripping branch because of the high resistance offered by the inductance of the upper branch. Therefore, it can be seen that the greater the rate at which the current rises, the lower the value of the total current at which the breaker trips. Thus, it is possible for the breaker to trip at the beginning of trouble rather than at the period where the trouble becomes fully developed. Due to the characteristics of the slow release relay, the circuit is allowed to come back to normal, and when the tension of the return spring is greater than the attraction of the electromagnet the circuit is once again closed. However, if the previous short circuit conditions still exist, the breaker is caused to trip again.

The type of circuit breaker used depends on the overload characteristics of the circuit, the supply source, the wire and insulation to be protected, and other factors. Air circuit breakers, molded-case circuit breakers, and oil circuit breakers are the types most commonly used. Thermal action, magnetic action and thermal-magnetic action are some of the tripping mechanisms utilized by circuit breakers to open a circuit.

4.9.3.1 Thermal Action Circuit Breaker

The tripping element of a thermal circuit breaker consists of a bimetallic strip and a heating element. The bimetallic strip is made up of two strips of metal with different coefficients of expansion. These two strips of metal are firmly attached together by welding or riveting, and one end of the resulting strip is fixed to the circuit breaker enclosure. The other end is fixed to the circuit breaker operating mechanism in such a way that if the strip bends it allows the mechanism to spring open. A heating element is connected in series with the load and is of such a size that it does not heat up to any extent except under overload conditions. When overload conditions exist the heating element causes the bimetallic strip to heat up causing the metals that comprise this strip to expand. Since these metals have different coefficients of expansion they tend to expand at different rates causing the bimetallic strip to bend in the direction of the metal with the lower coefficient of expansion. It is this bending action that trips the circuit breaker causing the circuit to open.

4.9.3.2 Magnetic Circuit Breaker Action

When the current of a heavy overload or short circuit energizes a magnet which has a predetermined setting the circuit breaker is instantly tripped. This is somewhat similar to the thermally operated circuit breaker described above except that a magnetic tripper is employed in place of the heater and bimetallic strip. Magnetic action, however, is not practical when used alone, except when it is used for some special applications. The magnetic tripping cannot be set low enough for safety and still handle a starting current. Magnetic action is found most valuable when used in combination with thermal action.

4.9.3.3 Thermal-Magnetic Action Circuit Breaker

Thermal-magnetic action combines the best features of both thermal action and magnetic action. The thermal elements protect on overload, where non-instantaneous tripping is desirable, such as in the case of a prolonged but not very great overload. The magnetic tripping element operates the circuit breaker instantly in case of a dangerous overload or short circuit.

4.9.3.4 Large Air Circuit Breakers

Large air circuit breakers consist of an operating mechanism, contacts, an arc interrupter, and usually a built-in overcurrent tripping device. These circuit breakers are characterized by their sturdy construction and high current carrying and interrupting ability. The tripping devices are also adjustable for operating time. They are particularly applicable for protection for larger loads over 200 amps or for smaller loads where the highest reliability of protection is desired or electrical actuation is required. Large air circuit breakers may be of the dual magnetic type or the thermal-magnetic type.

4.9.3.5 Molded-Case Circuit Breakers

Molded-case circuit breakers are smaller, less sturdily constructed, and do not have the electrical clearances that large air circuit breakers have. It is also difficult to actuate them electrically. Because of their small size and lower cost, they are used mainly for branch-circuit protection where the continuous current is less than 600 amps.

4.9.3.6 Oil Circuit Breakers

Oil circuit breakers are characterized by their simple construction and operation. They are operated by the "oil-blast" principle which is as follows: When the breaker begins to open, the arc between the contacts forms gases which set up a pressure that forces a stream of oil into the arc path, thereby introducing sufficient insulation to prevent re-establishment of the arc, thus interrupting the circuit. Oil circuit breakers may be used for loads up to 5000 amps.

4.9.4 Contactor

A contactor is a device for repeatedly establishing and interrupting an electric power circuit. Its function is to establish a power circuit, to carry the current during the time that the contactor is closed, and to interrupt the circuit. It may be manually operated, electrically operated or operated by an air cylinder. A contactor is normally designed to interrupt a circuit which is 10 times its full load rating: but it is not designed to interrupt a short-circuit current of high magnitude. Fuses and other high speed circuit breakers should be provided for such protection.

Magnetic contactors are the most important part of any automatic starter. A magnetic contactor is one which is actuated by electromagnetic means. In most large motors or generator circuits when automatic sequencing on or off of power is required the contactors are magnetically operated. For instance, a computer quite often requires an automatic sequencing on of its power supplies. First the filament supplies and after a suitable warming time delay the high voltage supplies are turned on. This is often accomplished by means of an automatic sequencing circuit operating magnetic contactor.

PART 3

CHAPTER 5

ELECTRONICS

5.1 INTRODUCTION

Electron tubes are of two basic types, vacuum or gas-filled. They vary from two element types to multi-element types. The elements found in vacuum tubes are cathodes, plates and grids. Electrical current conduction involving the movement of electrons from the cathode to the plate through a vacuum or gas-filled space is brought about by having negatively and positively charged electrodes. The movement of these electrons from cathode to anode may be compared to the movement of "ions" through a liquid in electrolytic conduction. However, the motion of the electrons, in a vacuum tube is always from the "hot" cathode to the "cold" plate. Electron tubes are replaced in many instances by solid state diodes such as germanium diodes and metallic-oxide rectifiers.

5.2 TUBE PARTS

5.2.1 Cathode

The cathode, one of the elements found in vacuum tubes, is an electrode which serves as a source of electrons. These electrons are emitted when the cathode is heated. This is known as thermionic emission. The number of electrons emitted from the cathode depends on the metallic substance used and the temperature of the cathode. The greater the temperature the greater the emission up to a maximum emission (saturation point), above which there will no longer be any increase. The saturation point depends on the type of material used for the cathode. Three of the substances commonly used for emitters are tungsten, thoriated tungsten, and oxide-coated metals.

Two types of heated cathodes are used in vacuum tubes. One is the filament or directly-heated type, which as its name implies, is heated to incandescence whereby electrons are then emitted. The other, the indirectly heated type, is one in which the cathode is an oxide-coated metal sleeve placed over a heating element. As an electric current is sent through the heating element, the temperature of the sleeve is raised sufficiently to cause electrons to be emitted.

It is also possible to secure combustion through a gas-filled tube with a cold cathode. In a cold cathode tube current will flow in either direction depending upon the relative potentials on the cathode and plate. However, the mechanics of emission is different from that of the heated cathode tubes. Gas filled, cold cathode tubes, depend upon an increase of ionization due to an increased potential difference between the plate and cathode. As ionization within the tube is increased, ionized particles bombard either the plate or cathode and liberate electrons. Thus a cold cathode tube may pass current in either direction.

5.2.2 Anode

The anode is the element in the electron tube which is usually maintained electrically positive with respect to the cathode. The anode is more commonly known as the plate. As a result of the potential difference between the plate and the cathode, a force is exerted on the electrons in the vacuum or gas-filled space. This force depends for the most part on the plate voltage. If the positive plate voltage is relatively low only a few electrons are drawn to the plate. As the plate voltage is increased more electrons from the cathode are drawn to the plate. The plate is usually a metal plate or cylinder which surrounds the cathode and grid.

5.2.3 Grids

The grid is an electrode that controls the number of electrons passing from the cathode to the plate. The effect of the grid is like that of a shutter which, opening and closing, controls the flow of electrons going through it from cathode to plate. This control is accomplished by changing the potential on the grid. When the grid is positively charged, it attracts electrons and by nature of its physical makeup (open wire mesh) and position (between plate and cathode), will allow the electrons to pass between the grid wires and go on to the plate, thereby increasing the plate current. When the grid has a negative potential with respect to the cathode it prevents the electrons from leaving the cathode, thereby decreasing the number of electrons reaching the plate. Therefore, when the grid is made negative with respect to the cathode, the plate current is decreased. Tubes may have one, two, three or more grids in the same envelope, each grid controlling tube conduction in a particular way.

5.2.4 Miscellaneous

Some specialized electron tubes, such as electron-multiplier phototubes and cathode ray tubes, have specialized parts which are discussed in the next section.

5.3 TUBE TYPES

Tube types will be discussed in this section under two separate categories, vacuum tubes and gas tubes.

5.3.1 Vacuum Tubes

In vacuum tubes, the only charged particles which move in the interelectrode space are electrons. These are negatively charged and therefore move from the cathode towards the plate.

5.3.1.1 Diode

The diode consists of two elements, a cathode and an anode or plate. When the cathode is heated sufficiently it emits electrons which are attracted towards the plate, providing that the plate is at a positive potential relative to the cathode. Therefore plate current will flow. However, if the plate is made negative relative to the cathode the electrons will be repelled by the plate and no plate current will flow.

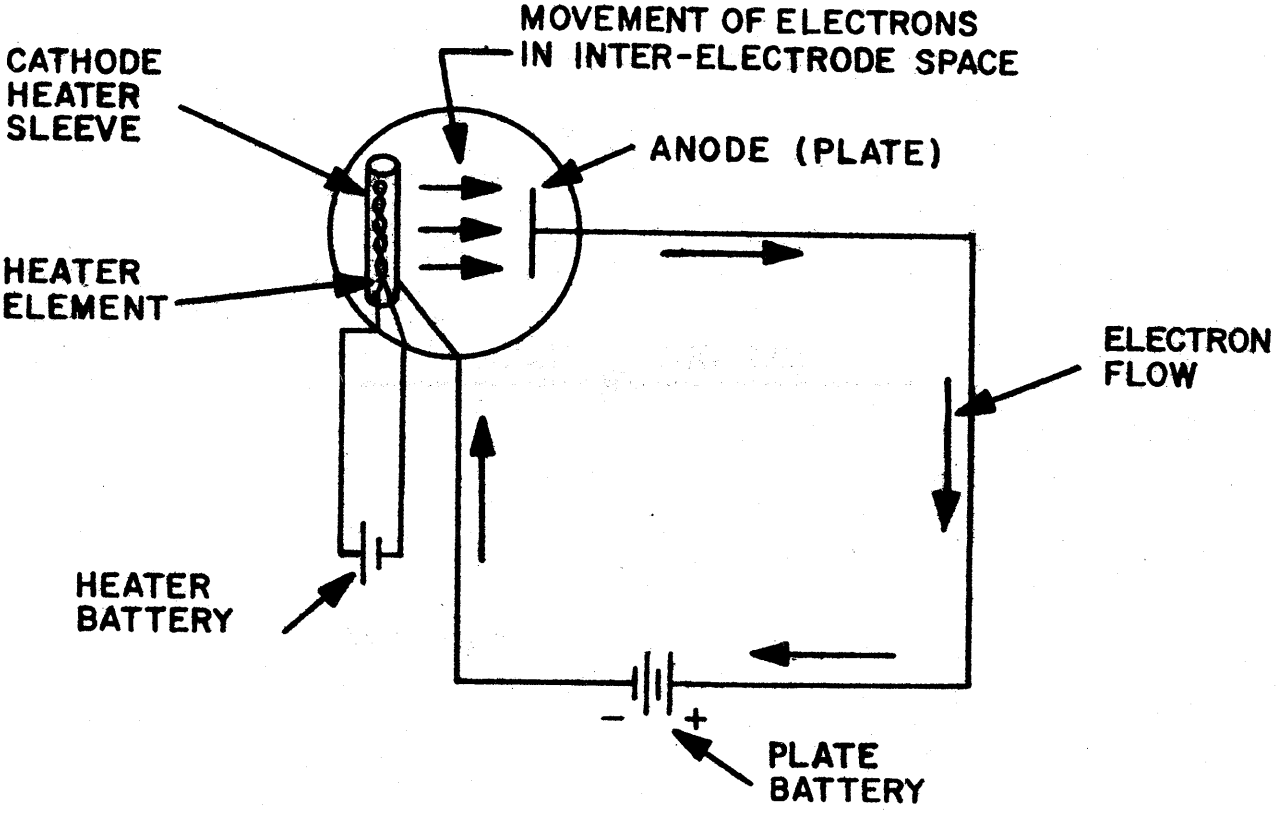

MOVEMENT OF ELECTRONS IN INTER-ELECTRODE SPACE
CATHODE HEATER SLEEVE
ANODE (PLATE)
HEATER ELEMENT
ELECTRON FLOW
HEATER BATTERY
−
+
PLATE BATTERY

Figure 3-76

Figure 3-77

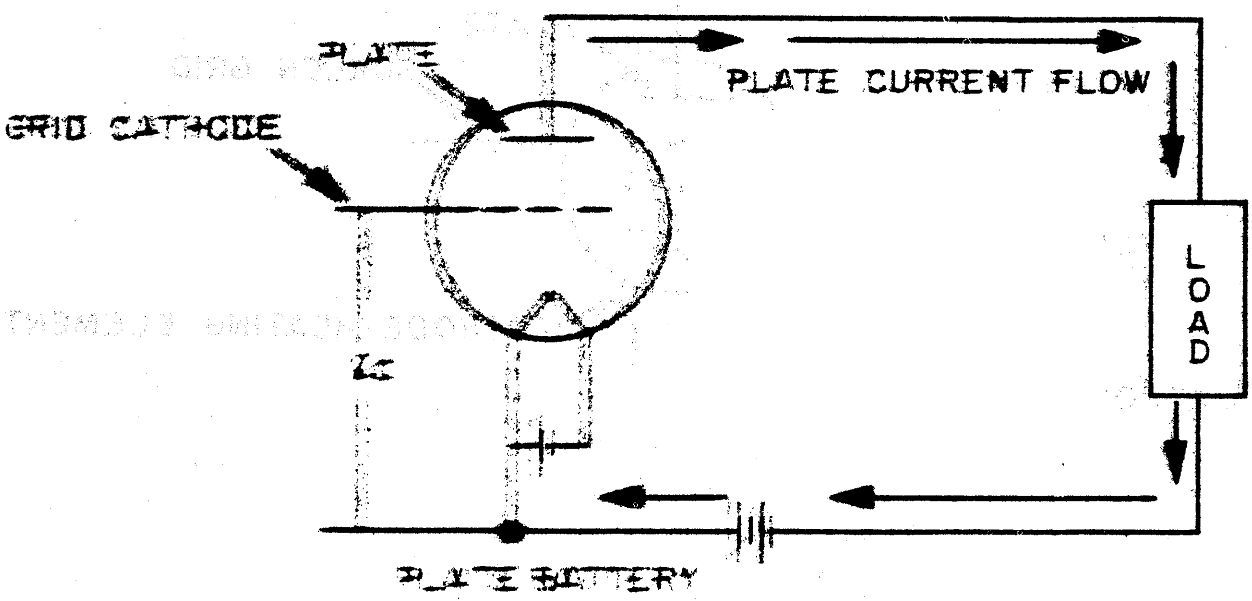


PLATE
GRID CATHODE
PLATE CURRENT FLOW
LOAD
Ec
PLATE BATTERY

Figure 3-78

By utilizing this characteristic of a diode, alternating current applied between the plate and cathode of this tube, would flow only during the time the plate was positive producing a direct current. This is known as rectification, and the device used as a rectifier.

Figure 3-76, shows electron current flow in a simple diode circuit. Figure 3-77, shows a simple schematic representation of a diode. Nothe that the triangle points in a direction opposite to that of electron flow, that is the triangle represents the plate and the straight line represents the cathode.

5.3.1.2 Triode

A triode is a tube having three elements: a plate, a cathode, and a grid. By varying the potential on the grid, which is located between the cathode and plate, the flow of electrons from the cathode to the plate can be increased or decreased. (Refer to section 6.2.3). The smallest negative voltage between grid and cathode which causes the tube to cease conducting is called the cut-off bias. A grid in a triode is called a control grid. Figure 3-78, shows a basic circuit diagram for a triode vacuum tube. The term e_c indicates the voltage applied between grid and cathode. The triode vacuum tube functions as an amplifier when a change of potential on the grid causes a change of plate current i_b which in turn causes variations of the voltage drop across the load in the plate circuit.

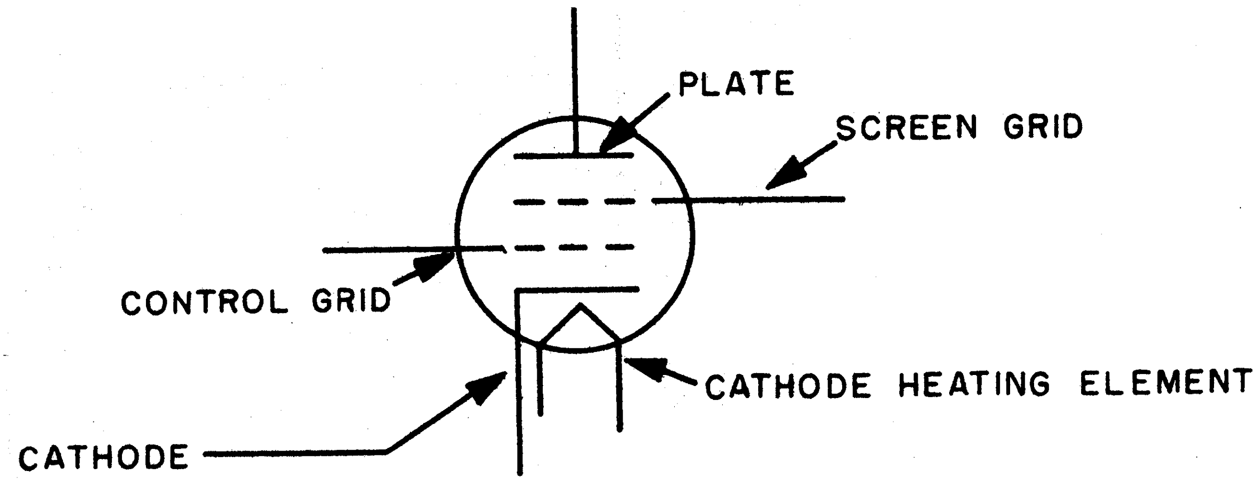


PLATE
SCREEN GRID
CONTROL GRID
CATHODE HEATING ELEMENT
CATHODE

Figure 3-79

5.3.1.3 Tetrode

A tetrode is a tube containing four elements: a cathode, two grids and a plate. The first grid or control grid functions as in a triode. Since the second grid forms an electrostatic shield or screen between the control grid and plate, it is called a screen grid. The screen grid is made of an open mesh network and is maintained at a fixed positive potential with regard to the cathode.

The addition of the screen grid reduces the interelectrode capacitance between the control grid and plate to a very low value. This interelectrode capacitance limits the effectiveness of a triode as a voltage amplifier at high frequencies because it constitutes a path for alternating current between the circuits connected to the control grid and plate. The electron current flow in a tetrode is influenced very little by plate voltage. It is mainly controlled by the control grid and screen grid voltages. Figure 3-79, is a schematic diagram of a tetrode.

5.3.1.4 Pentode

The pentode is a five element tube consisting of a cathode, a control grid, a screen grid, a suppressor grid and a plate.

The suppressor grid is incorporated into a tube in order to reduce secondary emission. In a tetrode, the high velocity of the electrons striking the plate, knock free electrons off the plate material into space. These secondary electrons are

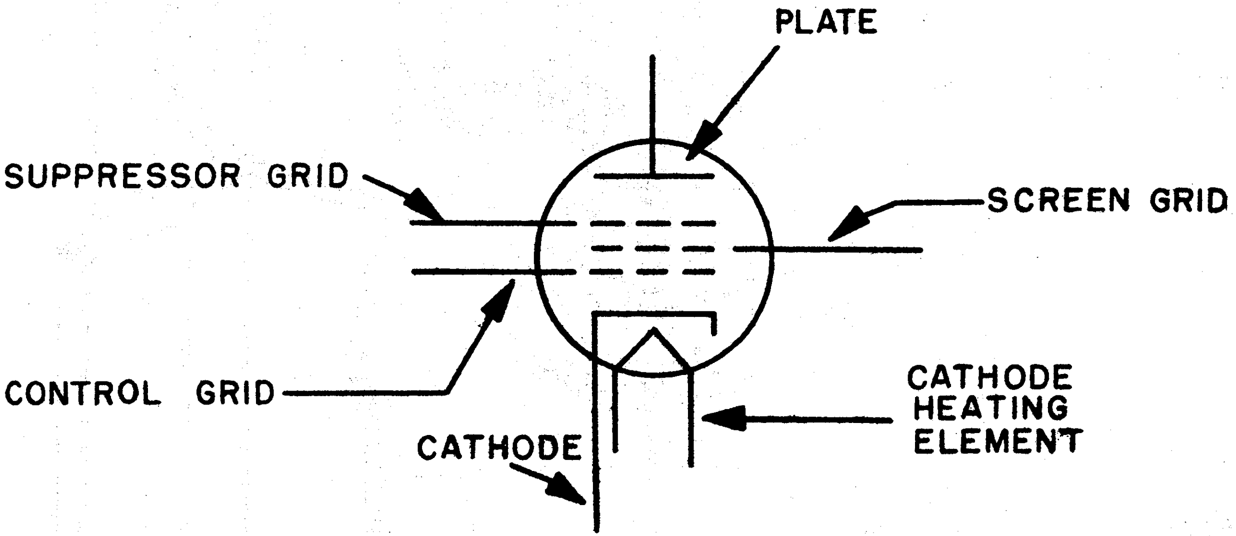


PLATE
SUPPRESSOR GRID
SCREEN GRID
CONTROL GRID
CATHODE
CATHODE HEATING ELEMENT

Figure 3-80

attracted to the screen grid, due to its positive potential. If any appreciable number of electrons produced by secondary emission is drawn over to the screen grid, the plate current is decreased and distortion occurs. The suppressor grid, ordinarily connected to the cathode, is placed between the plate and screen grid. This negative potential in the region near the plate drives the secondary electrons back to the plate and thus prevents them from reaching the screen grid. The suppressor grid is built with open construction so that it does not interfere with the normal passage of electrons from cathode to plate. For amplification purposes, the pentode has practically replaced the tetrode. Figure 3-80, shows a schematic diagram of a pentode.

5.3.1.5 Beam Power Tubes

A pentode when used for power amplification produces distortion when the plate voltages are lower than the screen voltage. That is, when plate voltage in a pentode becomes relatively low, the suppressor grid slows the volocity of electrons so that the plate acts as a cathode with respect to the screen. This results in an excessive density of electrons between the screen and plate grids.

A beam-power tube is a special type of tetrode in which electrons are confined by means of beam forming plates located on either side of the tube. In addition, the same spacing is used between wires on the control grid as on the screen grid.

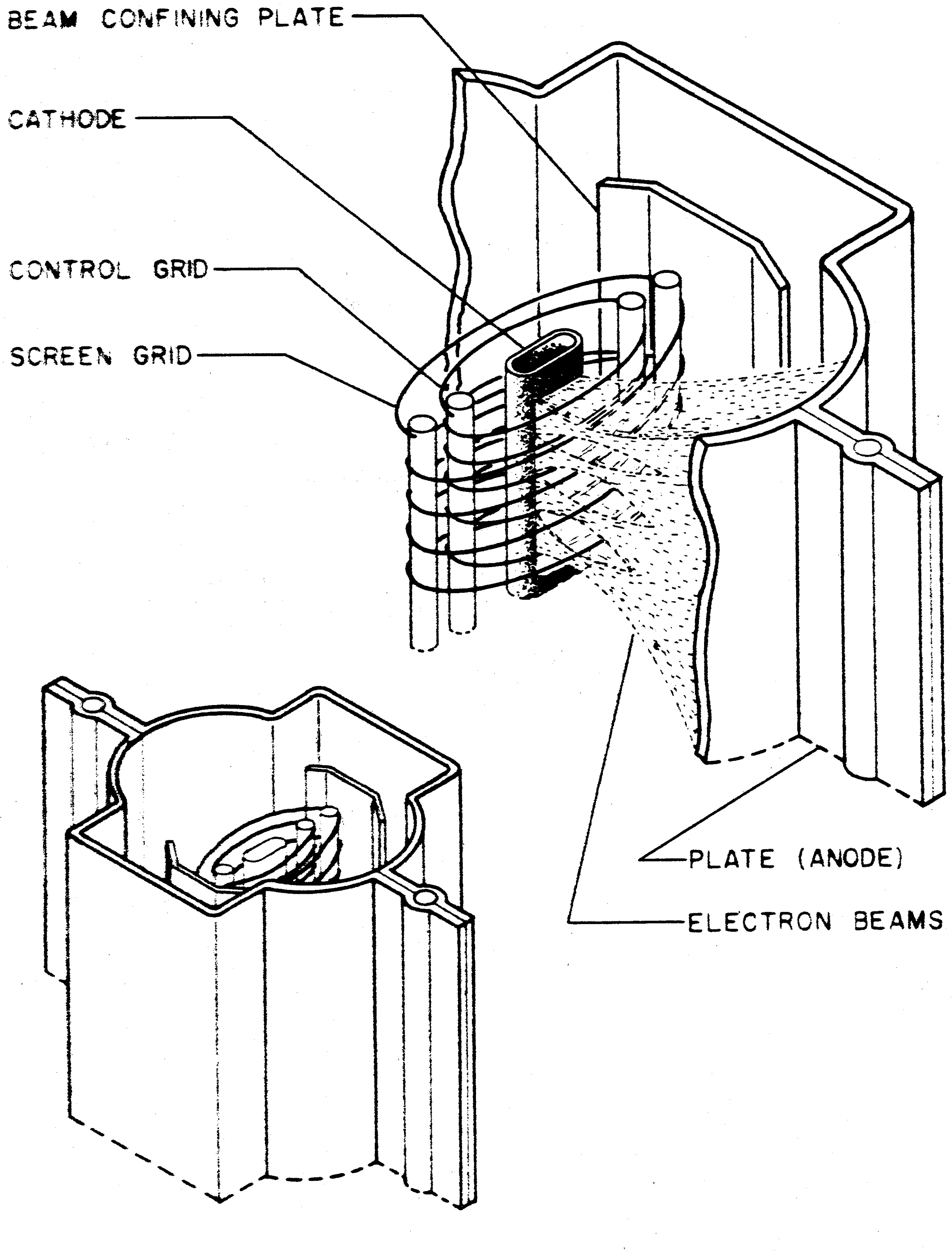

BEAM CONFINING PLATE
CATHODE
CONTROL GRID
SCREEN GRID
PLATE (ANODE)
ELECTRON BEAMS

Figure 3-81

The screen grid is also mounted so as to lie in the electronic "shadow" of the control grid. Consequently, an electron that passes through the control grid can proceed to the plate without striking the screen or be deflected appreciably by it. This type of construction also decreases the screen-grid current since no electron can strike the screen except by following a curved path around a curve of the control grid.

Figure 3-81, shows the internal structure of a beam-power tube and how electron beams are produced by the screen and control grids as well as the beam forming plates. The latter are connected electrically to the cathode and, in addition to confining the electrons to beams, **serve to** prevent any stray secondary-emission electrons from reaching the screen grid from the ends of the tube.

A beam-power tube when operated at the same voltages as a normal tetrode provides more power output for a given signal voltage on the control grid.

5.3.1.6 Multigrid Tubes

Vacuum tubes may be constructed with one cathode, one plate, and more than three grids. Such tubes are frequently used as frequency converters in which two different frequencies are mixed in order to obtain a third frequency.

In multigrid tubes having more than three grids, two of the grids serve as control grids. These control grids act to control the plate current in accordance with the combined effect of the separate voltages applied to each of the grids. The remaining grids are used as screen and suppressor grids.

Multigrid tubes are used widely in automatic-volume-control circuits and in volume expanders. In such cases, the amplification of a voltage applied to one control grid is changed by the voltage applied to the other.

Some examples of multigrid tubes are: a hexode containing a cathode, plate and four grids; a heptode with a cathode, plate and five grids; an octode with a cathode, plate and six grids.

5.3.1.7 Multi-Unit Tubes

In order to reduce the number of tubes in radio circuits, the electrodes of two or more tubes are placed within one envelope. There are various combinations of different type tubes, such as diodes, triodes and pentodes in one envelope. For example, a twin diode contains two diodes in one envelope; a twin triode contains two triodes in one envelope; a twin-diode-triode contains two diodes and one triode in the same envelope.

5.3.2 Gas Tubes, Thermionic Cathodes

Gas tubes are electron tubes in which a gas is deliberately inserted into the tube envelope. The gases commonly used are helium, neon, argon and mercury vapor. On schematic diagrams a small dot within the circle representing the envelope of a tube indicates it is a gas-filled tube.

In this section, the gas tubes discussed will be only those which have thermionic cathodes. Such cathodes emit electrons as a result of being heated. In later sections cold-cathode tubes and phototubes will be discussed.

5.3.2.1 Gas Diodes

In a gas-filled diode, the electrons flowing from the cathode to the plate encounter gas molecules. When an electron collides with a gas molecule, it results in one of the electrons of the molecule being detached. The electron thus produced, which is exactly like the one which came from the filament, becomes part of the electron stream traveling toward the plate. It may reach the plate directly, or it may strike another gas molecule and produce another free electron. This process continuously goes on in the space between cathode and plate.

The molecule of gas which has lost an electron is left with an excess positive charge and is therefore called a positive ion. The process of producing "ions" in a gas is called ionization. Positive ions travel towards the cathode instead of the plate. The current through the tube is therefore made up of two components: first, the electron stream flowing toward the plate; second, the stream of positive ions moving toward the cathode. These are both equivalent to a positive current from plate to cathode, so that the two components must be added to find the total current.

The behavior of a gas tube is the same as that of a vacuum tube until a voltage is reached called the ionization potential, striking potential, or firing point. It is at this point that the current in the gas tube increases. The ionization potential, which is the minimum voltage required before ionization can take place depends upon the gas used. The ionization potential for argon is 15.7 volts, and for mercury vapor 10.4 volts.

The current in a gas tube is higher than that in a vacuum tube at comparable values of operation. In vacuum tubes a cloud of electrons surrounds the cathode and prevents other electrons from leaving. These electrons are neutralized in a gas tube by the positive ions mixing with them. This allows a given plate voltage to draw more electrons away from the cathode, so that the current increases. There is also a negligible increase of current due to the additional number of electrons and positive ions produced by collision.

Once ionization has started, the action maintains itself at a voltage considerably lower than the firing point. A minimum voltage exists which is needed to maintain ionization. When the voltage across the tube falls below this minimum value, called the de-ionizing potential or extinction potential, the gas de-ionizes and the conduction stops.

Gas diodes can be used as electronic switches, which close at a certain voltage permitting current to flow, and then open at some lower voltage blocking the current flow. Such tubes have almost infinite resistance in circuits which have voltages too low for the tubes to operate, and very low resistances in the ionized state.

5.3.2.2 Thyratron

The thyratron is a gas-filled triode or tetrode in which a grid is used to control the firing potential. If the grid voltage is made sufficiently negative, the tube will not conduct current from cathode to plate, because the negative grid is more effective in driving electrons back to the cathode than the positive plate is in attracting them. If the value of the grid bias is gradually reduced, a point will be reached where there is sufficient attraction for the electrons to be drawn to the plate causing a small current flow. As the grid bias is further reduced, the electrons move toward the plate at higher speeds. When the average speed of the electrons reaches that corresponding to the ionization potential of the gas in the tube, ionization takes place and the current suddenly jumps. A current will also flow in the grid circuit because positive ions are attracted toward the grid as well as the cathode.

Once conduction starts in the plate circuit any increase of negative voltage on the grid has no effect on the plate current. It is possible to keep the tube from conducting by maintaining a sufficient negative bias on the grid, and conduction can be started by reducing the value of this negative bias. However, once conduction starts in the plate circuit the grid is no longer able to exert any control, even though it is made more negative than it was originally. The negative grid regains control only after the gas has been de-ionized by reducing the plate voltage to zero either by opening a switch in the plate circuit, or by using a plate voltage which is alternating and hence goes through zero twice in each cycle.

This loss of control by the grid is caused by the positive ions which are present due to the ionization of the gases. These positive ions migrate toward the grid wire under the action of the grid's negative potential, and cluster around the grid just as the negative electrons cluster around the cathode in a vacuum tube. The negative potential of the grid is therefore neutralized by these positive charges making it powerless to control the flow of electrons to the plate.

Thyratrons are useful in a wide variety of applications because of the relatively large amounts of power which can be controlled by means of the grid. For controlling purposes the thyratron depends upon the use of an alternating plate voltage which has a zero value twice in each cycle as opposed to the direct type of control found in vacuum tubes. In order to stop the current, the grid is made negative and as soon as the alternating plate voltage reaches zero the grid regains control. An important use of such a tube is in a sweep circuit for producing triangular waves.

The grid circuit of a thyratron may draw heavy current, especially if the grid becomes positive. In order to limit the grid current a resistor is placed in series with the grid.

5.3.3 Gas Tubes, Cold Cathodes

A gas tube using a cold cathode emits electrons from the cathode either by secondary emission or by field emission. In secondary emission, the electrons leave the cathode because other electrons strike it with sufficient force to knock these electrons from the cathode surface. In field emission electrons leave the cathode because of the presence of an external positively charged field at the surface. If this external field is sufficiently strong, it is capable of effectively pulling electrons from inside the cathode so that they break through the surface and escape into the surrounding space. Field-emission current densities of the order of thousands of amperes per square centimer are possible. Such enormous current densities can be maintained only for very short periods of time in order to avoid destruction of the cathode by the heat produced.

Secondary and field emission are not confined to gas tubes, but are used in such tubes frequently in lieu of thermionic emission.

Cold cathode gaseous tubes are used as rectifiers, voltage regulators, controls, counters, producers of light, current detectors, and radio frequency detectors.

The firing potential of cold cathode gas tubes is much higher than for thermionic or hot cathode gas tubes. Also, the tube tends to be erratic in that the firing point varies slightly during operation. The passage of current through the tube is indicated by a glow on the negative terminal. These tubes are often called glow-discharge tubes. Such tubes, when filled with neon gas, are called neon-glow tubes. When other gases are mixed with the neon, the color of the glow changes. When alternating current is applied to a neon glow lamp, both electrodes are surrounded with glow discharge.

5.3.4 Miscellaneous High-Frequency Tubes

In general, the tube types discussed up to this point function satisfactorily in the range from direct current to about 100 megacycles per second. Beyond this frequency, the effects of transit time, interelectrode capacitance, and inductance of the lead connections to the electrodes are important. These effects generally introduce undesirable characteristics.

Transit time refers to the time required for an electron to cross from cathode to plate. In a conventional tube dealing with a 1-megacycle wave, this transit time is of the order of 0.001 microsecond. Since the period of a 1-megacycle wave is one microsecond, the transit time is only $\frac{1}{1000}$ of the period. This is too small to change the characteristics of the wave from the input of the tube to the output. If the frequency were 1000 megacycles, the period would be 0.001 microsecond.

This is equal to the transit time. In this case a full cycle occurs in the time it takes for electrons to cross from cathode to plate. The variations in input voltages will not then be accurately reproduced in time in the plate circuit.

Interelectrode capacitances exist in a triode, for example, between grid and cathode, grid and plate, and cathode and plate. At very high frequencies, this results usually in undesirable capacitive coupling between the electrodes. Also, the lead connections to the electrodes, especially of the cathode lead, causes inductance effects, usually undesirable. In particular, the inductance of the cathode lead tends to resonate with the grid-to-cathode capacitance and to cause feedback coupling between the grid and plate circuits. This contributes to a loss at the input of the tube.

To overcome the defects mentioned above, various special tubes have been designed to be used at high frequencies. Miniature tubes using small interelectrode spacing, small electrodes, and high voltages overcome to some extent the defects introduced because of transit time and interelectrode capacitance. Lighthouse tubes are used to overcome the effects of lead inductance by making the connections of the electrodes parts of concentric transmission lines constituting the external circuit. Klystron tubes operate on the principle of velocity modulation of an electron beam. These tubes operate effectively as amplifiers and oscillators in frequency ranges extending upward from 600 megacycles per second. Magnetrons are used as oscillators with very high power output in ranges as high as 30,000 megacycles. Traveling-wave tubes are also used, primarily in the frequency range from 800 to 3600 megacycles. Klystrons, magnetrons, and traveling-wave tubes depend for their operation on the interaction of an electron beam with an lectromagnetic field.

5.3.5 Solid State Devices

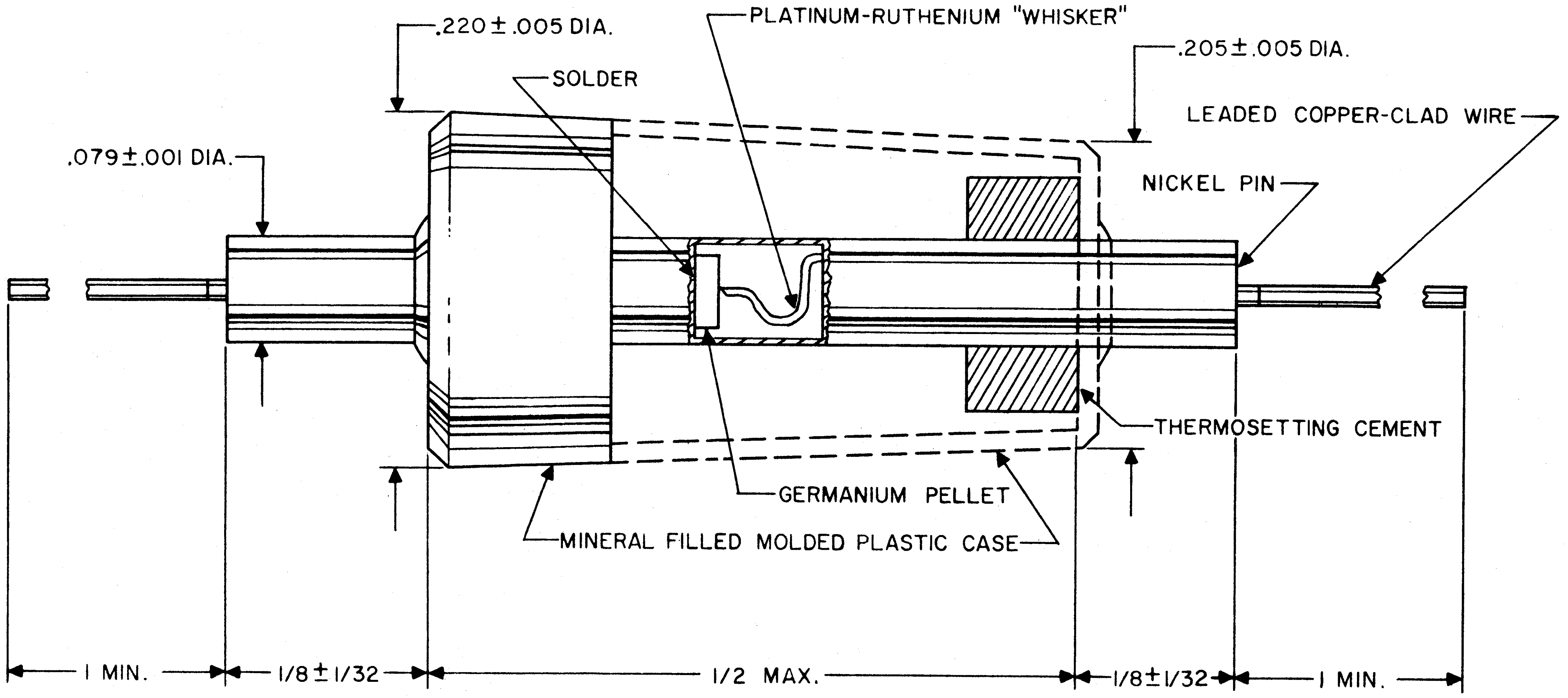
Solid state diodes such as semiconductors comprised of crystals and metallic-oxide rectifiers are used in many cases to replace vacuum tubes.

5.3.5.1 CRYSTAL SEMICONDUCTORS

Crystal semiconductors are composed of various materials that have electrical characteristics which may be regarded as intermediate between those of conductors and those of insulators. They are considered neither good insulators nor good conductors, under ordinary conditions, but can be made to exhibit some of the properties of each.

The crystals used most frequently in rectifiers are germanium and silicon semiconductors which do not have the abundant number of free electrons that are present in metallic conductors.

Crystal rectifiers may be formed by point or surface contacts between a metallic conductor, in the form of a wire, and a suitable semiconducting material such as germanium or silicon. The material of the metallic conductor is not of critical importance, but its resistance to corrosion is essen-



.220 ± .005 DIA.
PLATINUM-RUTHENIUM "WHISKER"
.205 ± .005 DIA.
SOLDER
LEADED COPPER-CLAD WIRE
.079 ± .001 DIA.
NICKEL PIN
THERMOSETTING CEMENT
GERMANIUM PELLET
MINERAL FILLED MOLDED PLASTIC CASE
1 MIN.
1/8 ± 1/32
1/2 MAX.
1/8 ± 1/32
1 MIN.

Figure 3-82

tial. The over-all dimensions and details of construction of a typical germanium rectifier are shown in Figure 3-82.

In silicon and germanium crystal rectifiers, the unidirectional conductivity or rectifying action takes place just under the point contact. The conducting agent moves in one direction only, due to a boundary layer which is formed in the crystal in this region. Electrons, in some cases are the conducting agents and in other cases, conduction is accomplished by the absence of electrons at certain energy levels. The absence of electron called an electron hole, behaves like a positive charge in supporting conduction. Applying an electric field to a piece of semiconductor material (placing a positive electrode on one end and a negative electrode on the other), causes electrons to drift through the crystal lattice toward the positive electrode and the holes towards the negative electrode. This constitutes a flow of current. Although electrons and holes move at the same time in opposite directions, the current component which each movement constitutes is in one direction.

In surface-contact germanium rectifiers the unidirectional conductivity results from a difference of impurities in the crystal. Rectification takes place at the boundary between two impurity states in the crystal. This type of rectifier is used mainly as a converter of a-c power to d-c, and is particularly useful for large current loads at low voltages and small current loads at high voltages.

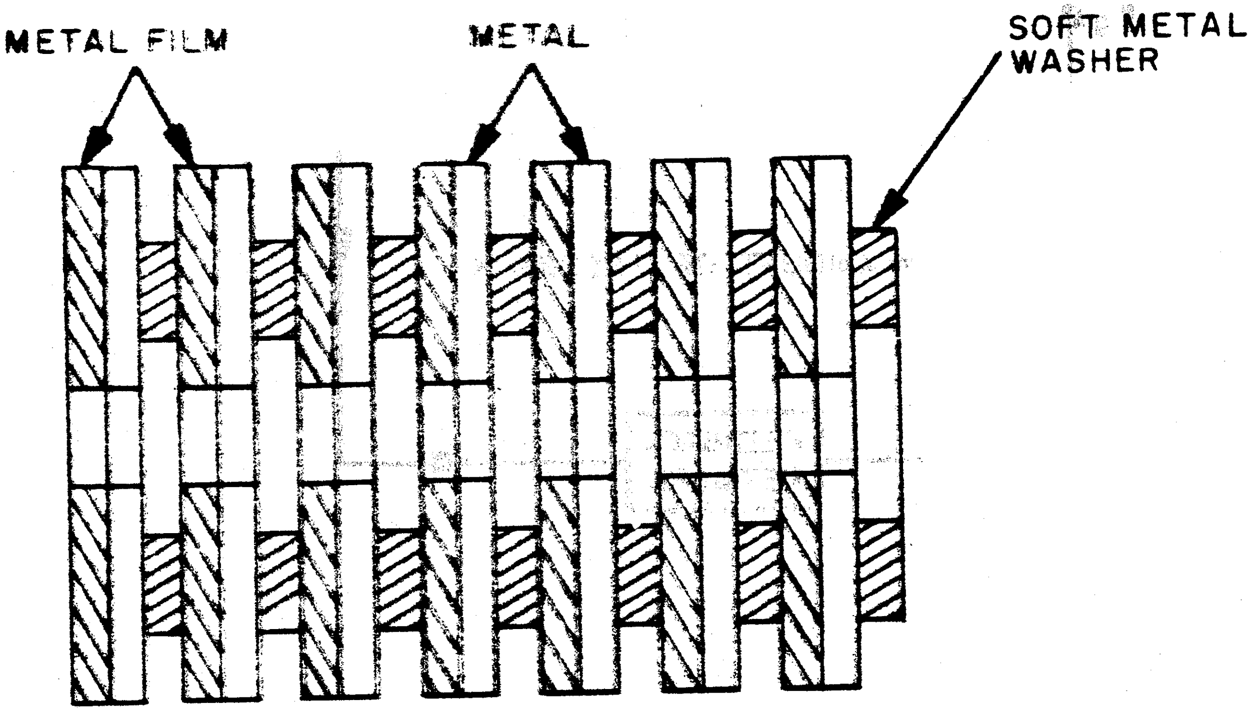


Figure 3-83

The function of the surrounding case of ceramic or plastic material of a crystal rectifier is to provide insulation, protection and accurate mechanical spacing. On one end of the rectifying diode is a band which indicates the cathode. If the unbanded end or anode is connected to the positive terminal of a battery maximum current will flow.

Crystal rectifiers are used in many circuits in place of diode vacuum tubes. Some of their characteristics are: high back resistance at comparatively high voltages, low capacitance (interelectrode and to ground), low forward resistance, no filament supply requirement, shorter electron transit time and more compactness and longer operating life than the diode vacuum tube.

5.3.5.2 Metallic-Oxide Rectifiers

There are combinations of thin films of metals that permit electrons to flow more easily in one direction than in the other. One combination used commercially is a thin film of copper oxide on a copper plate. The other combination is a thin film of selenium on a metallic surface such as iron or aluminum. Metallic-oxide rectifiers (see Figure 3-83) are used in many cases in place of vacuum tubes.

5.3.5.3 Copper-Oxide Rectifiers

In the case of copper-oxide rectifiers, electrons can pass from the copper to the copper oxide, much more easily than in the opposite direction. It is this ability to pass current in one direction that makes it useful as a rectifier.

The copper plate behaves like the cathode of a vacuum tube and the copper oxide as the plate. However, due to the thin layer of copper-oxide film the copper-oxide rectifier cannot withstand a very large voltage. Whenever moderate voltages are used, several layers of copper and copper oxide are built up into a stack. A washer of soft metal, generally lead is placed between the elements to provide uniform pressure between the copper and copper oxide. In order to increase the voltage at which the stack operates the individual cells may be connected series. The factor that determines the number of cells to be placed in series is the inverse voltage, for the rectifier circuit to be used, divided by the working voltage per cell. The current rating of the stack may be increased by connecting the cells of a stack in parallel. The total area of the cells which should be connected in parallel is determined by the average load current. In many cases, in order to get the desired voltage and current rating the cells of a stack are connected in a series-parallel arrangement. Several stacks can be connected in series or parallel arrangement for very high currents or voltages. A group of stacks is called an assembly.

Copper-oxide rectifiers, like other semiconductors have a negative temperature coefficient of resistance, so that the resistance to inverse voltage decreases with temperature. As the temperature rises the inverse current increases, which results in an internal power loss and temperature rise. Due to the increased power loss with temperature, the operating voltage must depend on the ambient temperature (30 degrees C average daily ambient temperature and a maximum ambient temperature of 40 degrees C) and the number of cells in a stack, since cooling is less effective with a long stack. The forward resistance increases with the life of the copper-oxide rectifier. This effect, which is called aging, is most rapid when the rectifier is new. The resistance becomes fairly stable after several months of operation if the rectifier is operating under normal conditions of temperature and load.

Copper-oxide rectifiers are used in relay and signal applications as well as in communication circuits.

5.3.5.4 Selenium Rectifiers

Selenium rectifiers are similar to the copper-oxide rectifiers in construction and operation. However, the semiconductor used on selenium rectifiers is a film of pure selenium on one side of a steel or aluminum plate. The use of aluminum reduces the weight of the plate. Selenium cells are assembled into stacks in a manner similar to that used for copper oxide cells.

The plates of the selenium rectifier provide support for the simiconducting material, serve as one electrode, and aid in transferring the heat losses to the air. The rectified current flow is from the metal backing plate through the selenium layer. The current rating of the selenium rectifier cells is about one-sixth ampere per square inch. If necessary it may be operated at a temperature of 55 degrees C. Insulating paint is often applied to the surface of the selenium cell in order to exclude moisture and protect the cell.

One of the most important advantages of the selenium rectifier is the high voltage at which it may be safely operated. The cells are usually rated for an applied voltage of 18 v rms, and 26 v cells are available. This is much above the voltage at which copper-oxide cells may be used. The higher forward resistance is compensated for by the higher operating voltage. Under optimum conditions the selenium rectifiers usually operate with a full-load efficiency of as much as 85 per cent.

The a-c supply voltage must be raised as much as 15 percent in selenium cells in order to maintain a constant d-c output voltage and current as the rectifier gets older. The forward resistance increases at a constant rate throughout the life of the rectifier. Another objectionable characteristic of the selenium rectifier is its decreasing resistance in the reverse direction when standing idle. If the selenium layer is not reformed, after a long period of idleness a large reverse current will flow when a voltage is impressed.

Selenium rectifiers are used for battery charging and for many control applications.

5.4 TUBE APPLICATIONS

5.4.1 Amplifiers

Amplifiers are devices which consist of one or more vacuum tubes and associated circuits. They are used to increase the strength of a signal. There are two general groups of amplifiers: voltage amplifiers and power amplifiers. Amplifiers are further classified according to the work they are intended to perform, as related to the operating conditions necessary to accomplish the purpose and the frequency range over which they operate.

5.4.1.1 Voltage Amplifiers

Voltage amplifiers are primarily concerned with the ratio of the alternating output voltage derived from the plate circuit to the alternating input voltage applied to the grid circuit. These output voltage variations across an impedance or resistance in the plate circuit are essentially of the same form as the input signal voltage impressed on the grid, but of increased amplitude. The plate current flowing through the load resistance causes the voltage drop which varies directly with the plate current. The ratio of the voltage variation produced in the load resistance to the input signal voltage is the voltage amplification, or gain, provided by the tube. Voltage amplification due to the tube is expressed by the following formula:

$$\text{Voltage amplification} = \frac{\text{amplification factor x load resistance}}{\text{load resistance + plate resistance}}$$

It can readily be seen that to obtain high gain in a voltage amplifier, a high value of load resistance should be used.

5.4.1.2 Power Amplifier

Power amplifiers are designed to deliver power to the load circuit while voltage amplification remains incidental. The ratio of output power to a-c power consumed in the grid circuit (driving power) is called the power amplification of the circuit. For power amplification, the load impedance is selected either to give maximum power with minimum distortion or to give a desired value of plate efficiency. Plate efficiency is determined by taking the ratio of output power to d-c input power to the plate (plate current times plate voltage). Amplifiers designed primarily for minimum distortion usually have a low plate efficiency, whereas plate efficiency may be quite high where distortion is permissible. Triodes, pentodes and beam power tubes are designed for power amplifier service. However, beam power tubes have a higher power sensistivity and efficiency and have higher power-output capability than the triode or conventional pentode types.

5.4.1.3 Amplifier Classification

Amplifiers are classified according to the purpose for which they are intended and according to the portion of the cycle during which plate current flows as controlled by the bias on the grid. There are basically four types of amplifiers: Class A, Class B, Class AB, and Class C.

5.4.1.4 Class "A" Amplifiers

A class "A" amplifier is one in which the grid bias and alternating grid voltages are such that plate current flows at all times.

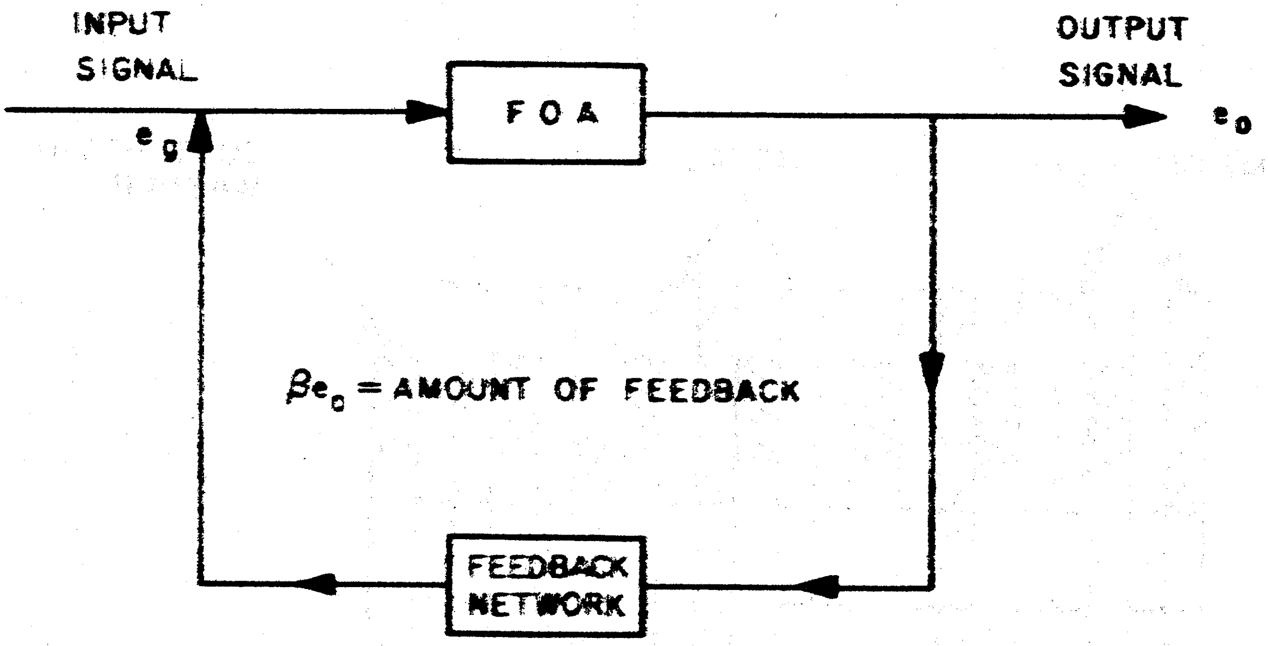

INPUT SIGNAL
e_g
F O A
OUTPUT SIGNAL
e_o
βe_o = AMOUNT OF FEEDBACK
FEEDBACK NETWORK

Figure 3-84

5.4.1.5 Class "AB" Amplifier

A class "AB" amplifier is one in which the grid bias and alternating grid voltages are such that plate current flows for appreciably more than half but less than the entire electrical cycle.

5.4.1.6 Class "B" Amplifier

A class "B" amplifier is one in which the grid bias is approximately equal to the cut-off value so that the plate current is approximately zero when no exciting grid voltage is applied. When an alternating grid voltage is applied, the plate current is a specific tube flows for approximately one-half of each cycle.

5.4.1.7 Class "C" Amplifier

A class "C" amplifier is one in which the grid bias is appreciably greater than the cut-off value so that the plate current in each tube is zero when no alternating grid voltage is applied. When an alternating grid voltage is applied the plate current flows for appreciably less than one-half of each cycle.

5.4.1.8 Feedback Amplifier

Feedback is a term which is applied to the process of transferring energy from the output circuit of a device to its input circuit. There are two types of feedback: regenerative or positive feedback and degenerative or negative feedback. A feedback system is shown in Figure 3-84. Part of the amplifier output voltage e_o is fed back to the amplifier input. The gain of this amplifier is

$$\text{Gain} - \frac{\mu}{1 - \mu\beta}$$

Where μ (mu) is the gain of the amplifier without feedback, and β (beta) is the portion of the amplifier output voltage used as feedback.

5.4.1.9 Positive Feedback

Regenerative, or positive feedback, are the terms used when the impulses fed back are in phase with the input, thereby increasing the initial input. This occurs when the quantity $1 - \mu\beta$ of the formula in section 2.4.1.8 is less than 1, increasing the gain of the amplifier. However, by increasing the gain of an amplifier an undesirable distortion or noise (that is introduced by the amplifier itself) is amplified and exaggerated. Whenever a distortionless output is desired, positive feedback is not used.

The amplifier becomes an oscillator when the quantity $\mu\beta$ is increased until it equals 1, causing the regeneration to become so large that sufficient energy is fed back to maintain the operation of the system indefinitely.

5.4.1.10 Negative Feedback

Degenerative, or negative feedback, are the terms used when the impulse fed back is 180 degrees out of phase with the initial input, thereby reducing the actual input to the amplifier. This negative feedback occurs when the quantity $1 - \mu\beta$ is greater than 1, decreasing the gain of the amplifier. Usually in negative feedback, $\mu\beta$ is made much larger than 1, such that $1 - \mu\beta$ may be considered equal to $-\beta$. The gain of the amplifier is then small, and may be expressed as

$$\text{Gain} = -\frac{1}{\beta}$$

Some of the advantages of negative feedback are: (1) Improving the waveform at the output of an amplifier by reducing the distortion which is introduced "within" the amplifier. The amplitude distortion of the input signal caused by the nonlinear characteristic of the amplifier is generated within the tube and therefore is not amplified in the output. Consequently by feeding back a fraction of the output, out of phase with the input, the distortion component of the degenerative voltage is amplified by the same factor as the input signal. As a result, the amplitude of the desired signal is considerably reduced and there is practically no distortion in the net output. However, there will be no effect by negative feedback, on amplitude distortion caused by the flow of grid current in the input stage of an amplifier. Since this distortion occurs at the grid of the amplifier it will be amplified in the same ratio as the desired signal; (2) Reducing the noise introduced within an amplifier is accomplished in the same manner that nonlinear distortion is reduced. However, if the noise component is present in the input signal, negative feedback will have no effect. It can only decrease the noise which is generated within the section of the amplifier around which the feedback takes place; (3) Amplifiers may be made to have any desired frequency response by designing a feedback network in which β

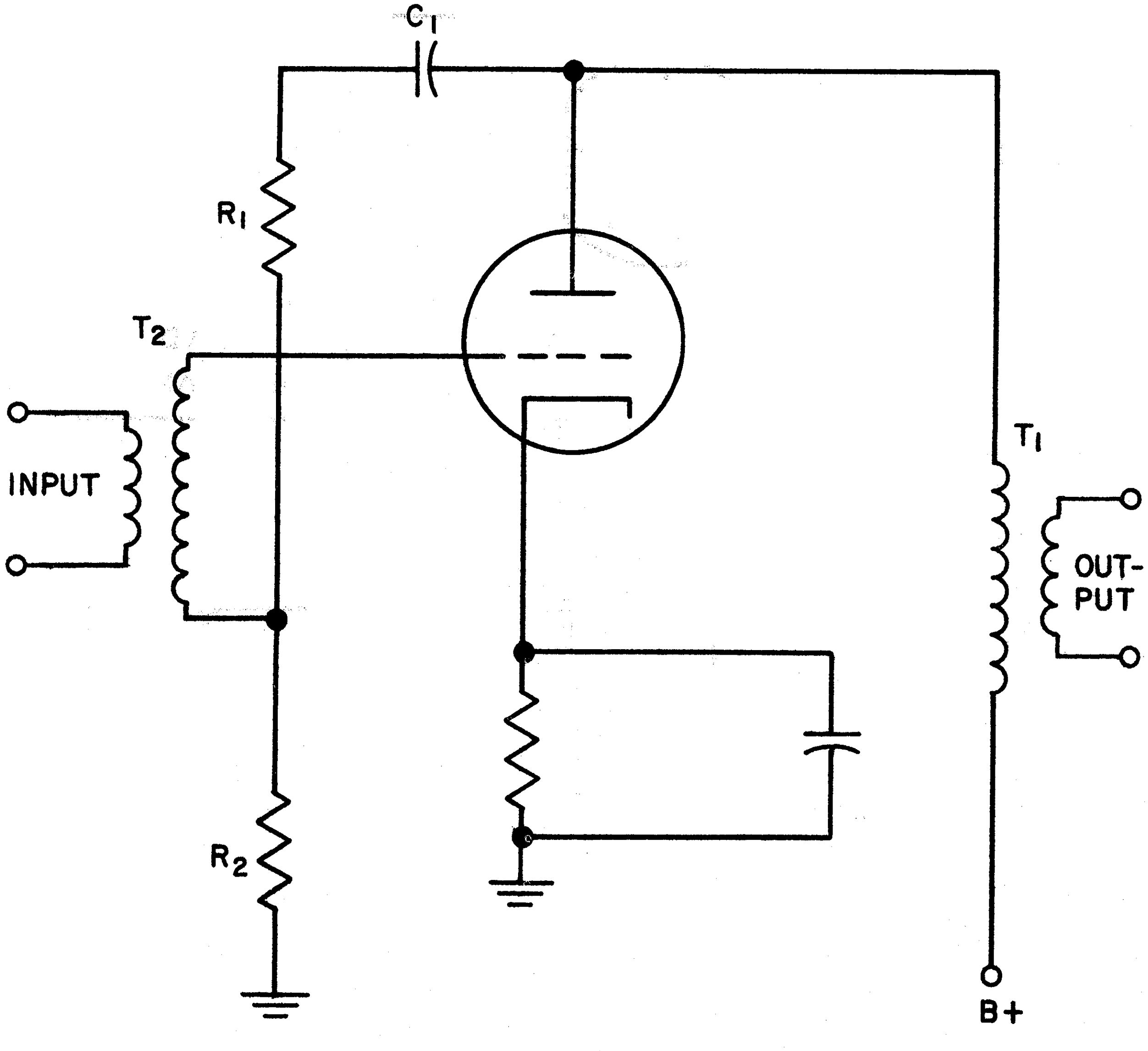


Figure 3-85

has the proper characteristics. This can be done because the gain of a degenerative amplifier is equal to $-\frac{1}{\beta}$. The frequency response of an amplifier can be improved, if feedback is obtained through a resistive network where the feedback factor β is independent of frequency; (4) The gain of the amplifier can be made independent of the load impedance, if the load impedance does not form part of the feedback path. If an amplifier is feeding a certain load resistance which is reduced by connecting other loads in parallel with it the output voltage tends to drop. Under these conditions, if degenerative feedback is provided, the feedback also tends to decrease, causing the effective gain to increase. This increase in gain maintains a nearly constant voltage output, by offsetting the tendency of the output voltage to drop. Any increase in the voltage output caused by the load being reduced or removed is checked by increased feedback; (5) It is apparent, since the gain of a degenerative amplifier is equal to $-\frac{1}{\beta}$ and is proportional only to the feedback factor that the effects of variations of battery voltage or aging of tubes are eliminated. Therefore, if the feedback factor is large the gain stability of an amplifier is improved by negative feedback.

There are several methods generally used to obtain degenerative feedback. Some of the methods are: (1) The voltage feedback method which is obtained by the use of a voltage-divider circuit, Figure 3-85. An input signal makes the grid of the tube less negative, resulting in an increase

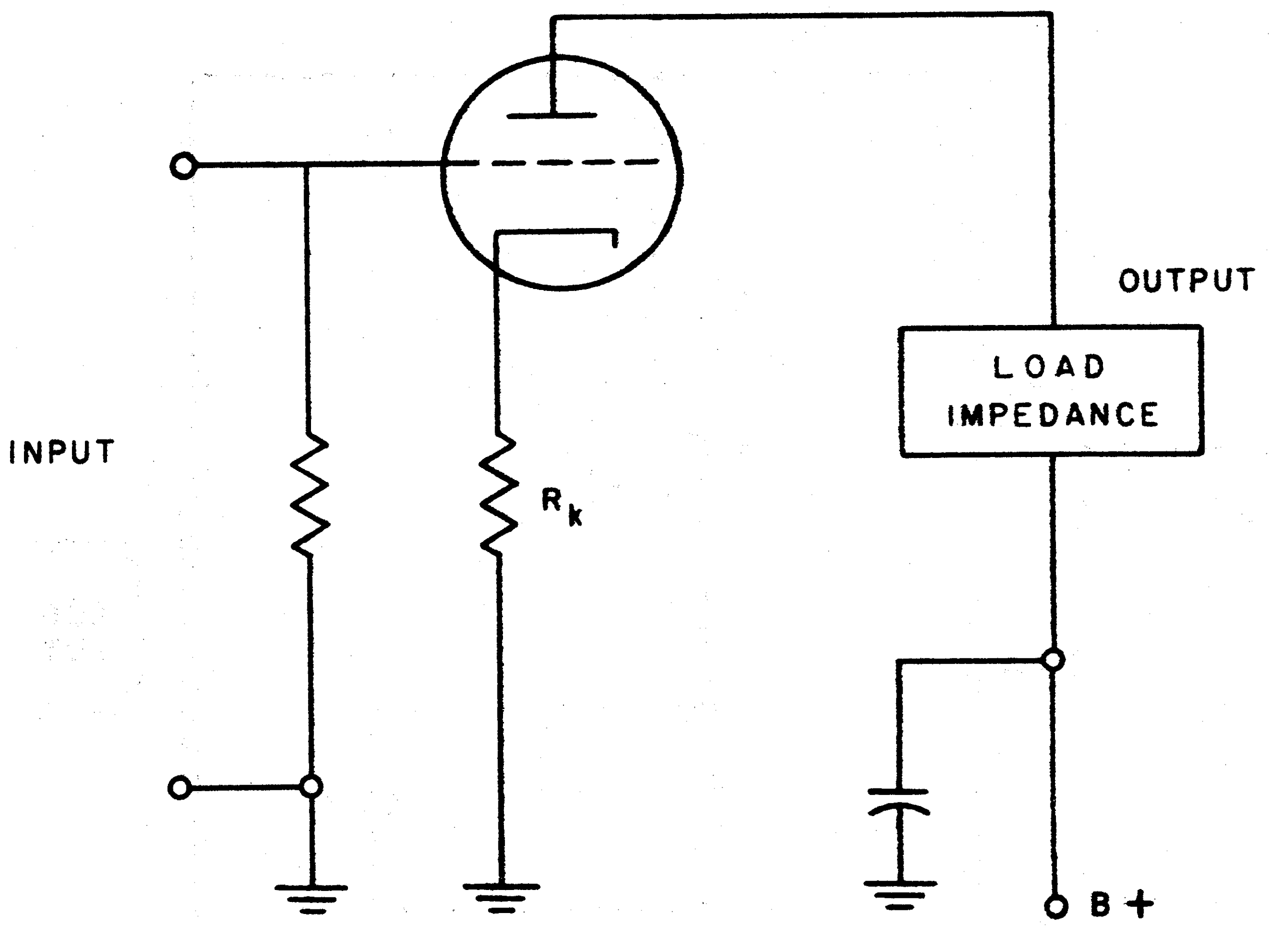


Figure 3-86

in plate current which in turn decreases the potential at the top of the primary of the output transformer T_1. This drop of potential is transmitted through the circuit elements; capacitor C_1, resistor R_1, and resistor R_2. The decrease of potential at the top of R_2 causes a negative-going voltage to be added to the positive-going signal voltage on the grid by way of the secondary of the input transformer T_2; (2) The current feedback method illustrated in Figure 3-86 is another method employed with degenerative feedback. The flow of plate current through the unbypassed cathode resistor R_K develops a feedback voltage. However, since this resistor is located between the cathode and the grid of the tube, the voltage developed across it appears in series with the input signal of the tube. The plate current and the voltage drop across R_K increase when a positive signal appears on the grid of the tube. As the voltage increases across the cathode resistor the grid becomes more negative relative to the cathode, which is the reverse of the signal voltage.

5.4.1.11 Amplifier Frequency Ranges

Vacuum-tube amplifiers are classified according to the frequency range over which they operate. These are the audio-frequency amplifiers which may be transformer-coupled; video-frequency amplifiers which are usually resistance-coupled amplifiers, requiring flat gain characteristics over a very wide frequency range; intermediate and radio-frequency amplifiers which are designed ordinarily for tuned- circuit coupling, although in actual operation they may resemble

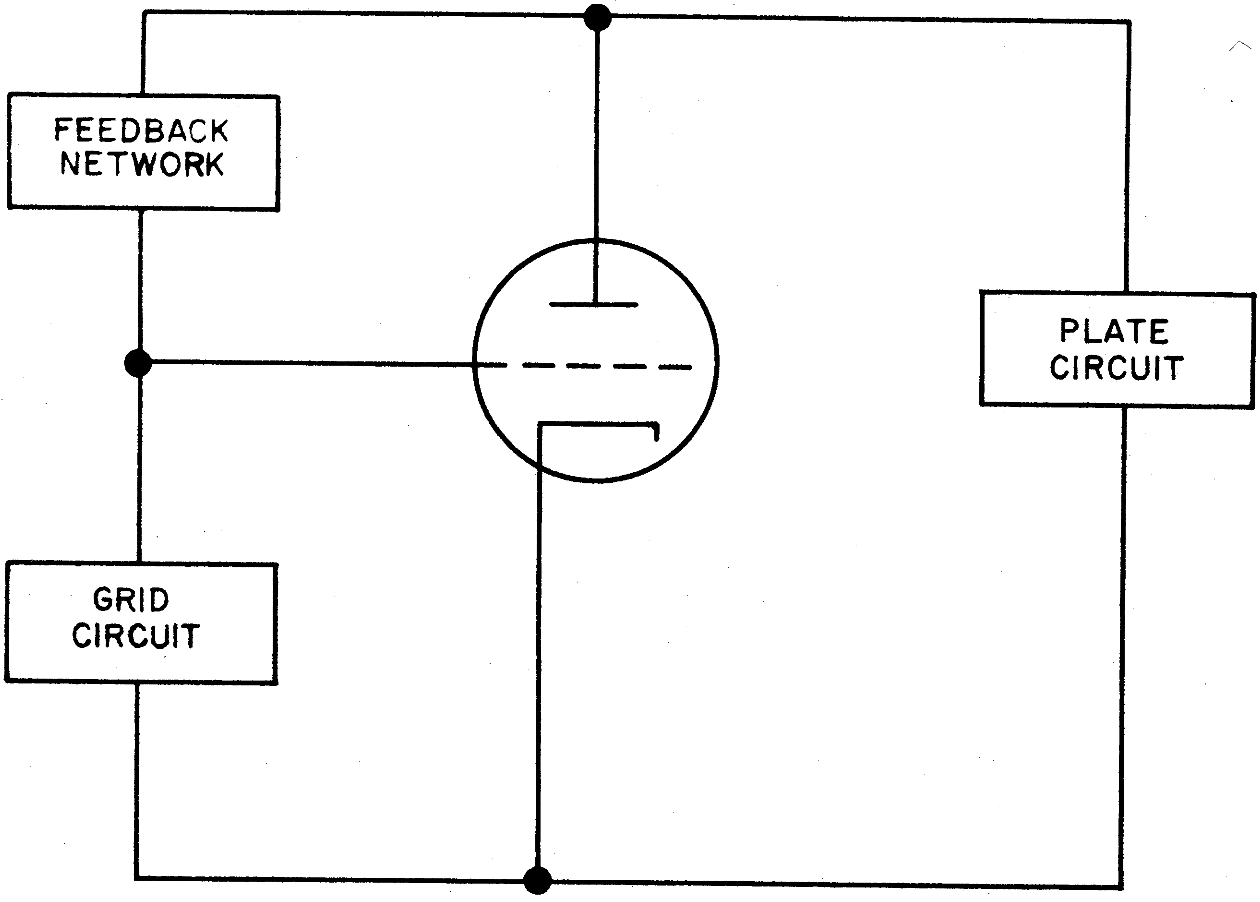


Figure 3-87

either the transformer-coupled or the impedance-coupled circuit.

5.4.2 Oscillators

Vacuum tubes used as ocillators are essentially energy converters which change d-c electrical energy from the plate-circuit power supply into a-c electrical energy in the output circuit. This is accomplished by the use of the amplifying ability of the vacuum tube in such a manner that the tube generates sustained oscillations. When the tube acts as an amplifier the energy is increased from the grid circuit to the plate circuit. Part of this plate-circuit energy may be fed back to the grid and used to supply the input power (see Figure 3-87). By supplying its own input, the tube is able to oscillate at a frequency determined by the constants of the circuit. Any small voltage change in the plate or grid circuit can be transferred from one to the other causing the tube to oscillate. Due to the amplifying ability of the tube, the amplitude of the signal is further increased. The plate voltage of the tube increases and decreases alternately during the operation of the circuit. A maximum value of voltage variation determinded by the operating characteristics of the tube, soon is reached. The tube now is said to be oscillating with maximum power being developed. The varying plate current, caused by the oscillator, flows through the load impedance in the plate circuit developing an alternating voltage across this load.

5.4.3 Voltage Regulator Tube

Voltage regulators are devices used with power supplies to maintain constant output voltages regardless of the large changes of load current drawn from the power supply or changes in the input voltages. Electronic voltage regulators are mainly used with rectifier power supplies.

5.4.3.1 Glow-Tube Regulator

In a neon-glow tube regulator (glow-discharge tube) the voltage across the tube remains constant over a fairly wide range of current through the tube. This is due to the degree of ionization of the gas in the tube, which varies with the amount of current that the tube conducts. When a large current is passed, the gas is highly ionized and the internal impedance of the tube is low. When a small current is passed, the gas is lightly ionized and the internal impedance of the tube is high. The product (IR) of the current through the tube and the internal impedance of the tube is practically constant over the operating range of the tube. The glow tube is placed in parallel with the load.

5.4.3.2 Vacuum Tube Regulator

Vacuum tubes may be considered as variable resistors. The d-c plate resistance, R_p is simply the plate-to-cathode voltage divided by the current through the tube, when it is conducting a direct current. For any given plate voltage the value of R_p depends upon the current through the tube, which in turn depends upon the grid bias.

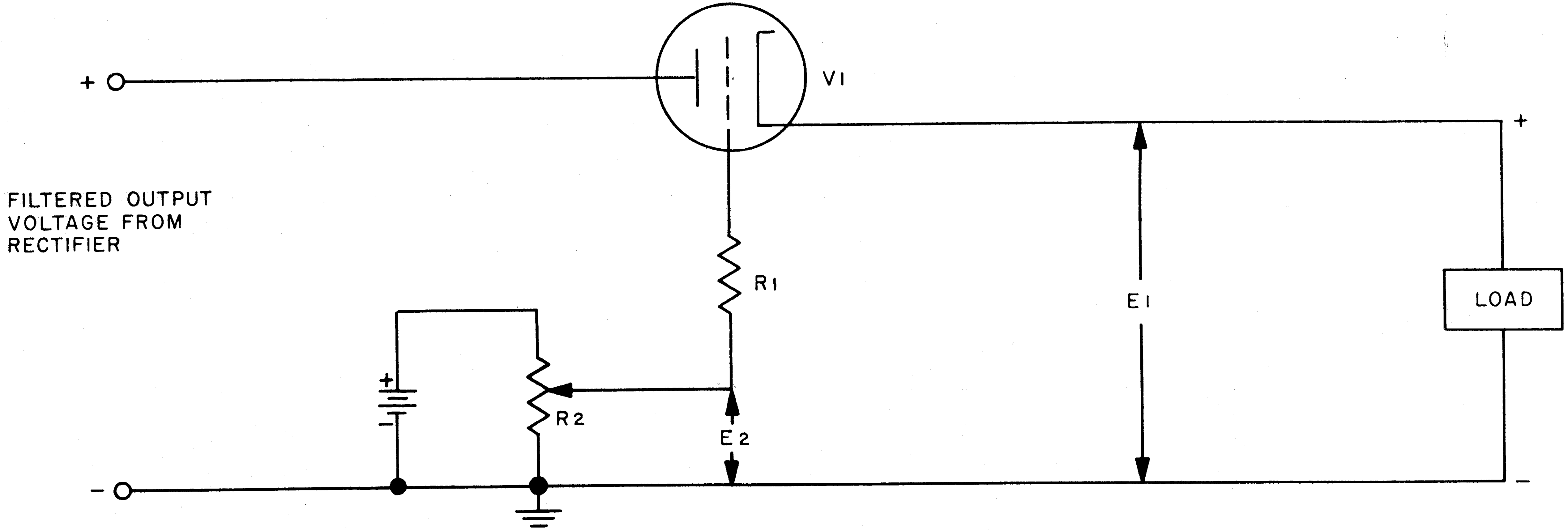


Figure 3-88

The vacuum tube regulator is used in the following manner:

Assuming that the voltage across the load is at the desired level, the cathode is positive relative to ground by a voltage E_1 (see Figure 3-88). The grid can be made positive relative to ground by a voltage E_2 which is less than E_1. R_2, a potentiometer is adjusted until the bias, which is equal to E_1-E_2, is sufficient to allow V_1 to pass a current equal to the load current. At this condition, the resistance of V_1 is established at the proper value to reduce the rectifier output voltage to the desired load voltage.

The voltage at the cathode of V_1 will increase if the rectifier output voltage increases. However, as E_1 increases the bias on the tube increases and the effective plate resistance of the tube becomes greater. Therefore the voltage drop across V_1 is greater. The increased voltage drop across V_1 should be approximately equal to the increase of voltage at the input to the regulator. The load voltage should therefore remain essentially constant. The resistor R_1 is used, in this circuit, to limit the grid current. This is necessary because the battery is not disconnected when the power is turned off.

The amplification of tube V_1 permits operation of the regulator on small variations of the load voltage, whether these variations are caused by supply-voltage fluctuations or changes in the load. In Figure 3-88 if the load resis-

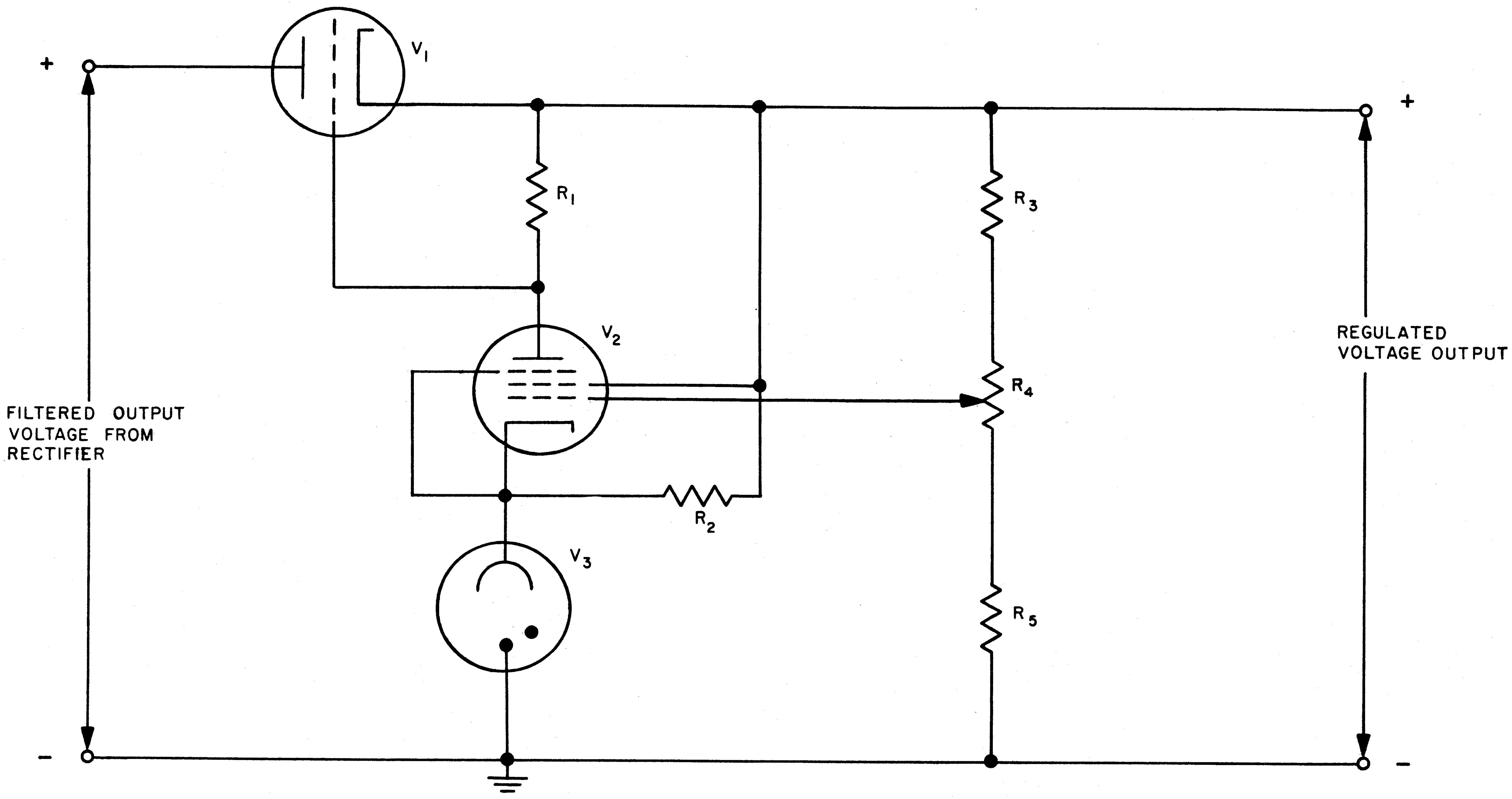

+
V1
+
R1
R3
V2
REGULATED VOLTAGE OUTPUT
R4
FILTERED OUTPUT VOLTAGE FROM RECTIFIER
R2
V3
R5
−
−

Figure 3-89

tance decreases, the load current increases, lowering the voltage E_1 by the increases voltage drop in V_1. The bias on V_1 decreases, since it is made up of E_1-E_2 and the cathode of V_1 becomes less positive. The resistance of V_1 lowers, due to the decrease in bias, causing the load voltage to rise to normal. The end result is that the total resistance across the power source is less, but the voltage division is so altered as to provide the desired voltage across the load.

Another regulator, Figure 3-89 uses a triode, a glow tube (fixed bias) and a pentode vacuum tube (high apmplification) is a degenerative type regulator. It is used to produce an output voltage independent of fluctuations in the a-c supply and changes in the load current over a wide range. It is used in the following manner: The output voltage is developed across the bleeder resistor R_3, R_4, and R_5 in parallel with the load resistance. The total voltage divider of the circuit is composed of two parts: the bleeder resistor R_3, R_4 and R_5 and the plate to cathode resistance of vacuum tube V_1 through which all the load current must flow. The remaining elements in the circuit are used to control the resistance of V_1 and thereby to maintain a constant voltage across the load.

The regulated voltage output of the regulator acts as the plate voltage of V_2 while the glow tube V_3 maintains the cathode of V_2 at a constant positive potential. The grid potential of V_2 is set by the potentioneter, R_4 so that it is less positive than the cathode by an amount (the bias) which causes V_2 to conduct. The voltage drop across R_1, caused by the plate current through it, is the bias on tube V_1.

The normal resistance of V_1, which is controlled by adjusting potentiometer R_4 is used to set the value of the load voltage which the regulator is to maintain.

When the load voltage tends to rise, either from a decrease in the load current or from an increase in the input voltage, the voltage on the grid of V_2 tends to rise (become less negative). The cathode voltage remains practically constant. Due to the decrease in bias, V_2 conducts more current resulting in a larger voltage drop across R_1, which is the bias voltage for V_1. The increased bias (more negative potential) on the grid of V_1 causes an increase in plate resistance. This causes a larger portion of the available voltage to appear across V_1 keeping the load voltage practically constant. If the load voltage tends to fall a similar action will occur.

Since all of the load current must pass through V_1, the tube must be capable of passing a large current. In some regulators several identical tubes may be put in parallel in order to pass the required current.

The degenerative type regulator shown in Figure 3-89 is used to stablize the output voltage of rectifier power supplies. This regulator is very effective in removing ripple from the output of rectifier power supplies, due to its sensitivity to small changes of input voltage. It also helps to filter the output of a rectifier, even though filters usually are used in connection with regulators.

5.5 RATING AND TOLERANCE

Ratings of electron tubes are based generally on two considerations. First, design-center maximum ratings are variations expected and permissible because of their arising from usual fluctuations of the voltage of alternating or direct current lines, storage batteries, or other sources of power. Second, absolute maximum ratings are those caused by the physical and chemical characteristics of the various materials from which tubes are made. Included here are ratings due for example, to the ability of tube elements to dissipate heat, to avoid arc-over, to emit electrons when heated, to avoid the evolution of gas in the tube envelope, and to avoid softening of a tube glass envelope. Various factors considered in tube ratings are listed below.

5.5.1 Heat Dissipation Within A Tube

The inside of a tube heats up for several reasons. One important reason is the heating of a cathode in a thermionic tube. Another reason is the heating of the plate due to the electron bombardment from the cathode. This heat must leave the tube through the tube walls. In the process, the tight seal may be weakened by softening of the glass wall, if glass is used, thus weakening the envelope. Gas evolution may also take place. Because of this limitation, glass walls are generally not used in tubes having more than one kilowatt rating.

5.5.2 Plate Heat-Radiating Ability

The plate is heated both by electron bombardment from the cathode and heat energy radiated from a thermionic cathode. The heat-radiating ability of the plate is a limiting factor in the tube rating. In some tubes, the melting point of the plate must be considered.

5.5.3 Electron Emission From The Plate

Electrons are emitted from the plate either because it becomes hot or because of secondary emission caused by electrons striking the plate. This electron emission is undesirable and becomes intolerable when it exceeds certain limits depending on the use of the tube. In a diode rectifier, for example, electron emission from the plate becomes reverse current conduction when the plate voltage polarity reverses as it does every half cycle in the application of an alternating current voltage.

5.5.4 Cooling Of The Plate

To remove some of the undesirable features mentioned in 5.5.1 through 5.5.3, the plate may be made of material having great heat radiation emissivity, such as graphite or nickel coated with carbon. Tantalum as a plate tends to absorb gas and thus improves the vacuum inside a tube. Large electron tubes may be constructed with water-cooled plates. Forced air circulation is common.

5.5.5 Maximum Voltages Between Electrodes

Maximum voltages between electrodes is limited by the danger of arc-over or flashover. This is particularly critical when electrode leads are near each other. For the usual receiver-type tubes in which leads are only one or two millimeters apart, the maximum potential difference between leads is of the order of 500 to 1,000 volts.

5.5.6 Maximum Peak Inverse Voltage

In rectifier circuits, the plate during each half cycle of an applied alternating current voltage becomes negative with respect to the cathode. Tubes are designed to withstand an inverse-peak plate voltage of one to two times the peak value of the applied alternating voltage. This is to avoid breakdown. A safe maximum value for gas-filled tubes is required to prevent arc-back in the tube when a reverse voltage is applied. In addition, cathode and plate leads in all tubes must be sufficiently separate to prevent a spark between them. Tubes are sometimes immersed in oil to enable them to withstand greater inverse peak voltages without flashover.

5.5.7 Peak Heater-Cathode Voltage

In tubes having separately heated cathodes, the potential difference between the cathode sleeve and the filament heater inside it cannot consistently exceed more than certain highest instantaneous values.

5.5.8 Grid Considerations

Grids can become heated as a result of heat energy radiated from the cathode or plate. This may cause electron emission from the grid by thermionic means. Also, electrons striking a grid may cause secondary emission. Such electron emission may lead to undesirable grid current, especially when the grid is more negative than, for example, a neighboring cathode. To avoid this, the temperature at which tubes are permitted to operate must be low, heat radiation from the grids by fins must be provided, or heat must be conducted away from the grid by good heat conductors. As some tubes age, grid emission current is aggravated because the oxide-coated materials from the cathode, as they evaporate, adhere to the grids. This increases the emissivity of electrons from the grid.

5.5.9 Maximum Peak Plate Current

Every tube can safely carry certain peak instantaneous plate currents in the direction of normal electron flow. Currents greater than these peak plate currents, occurring repeatedly, will shorten the life of the tube or damage it.

5.5.10 Maximum Direct Current Output

Every rectifier tube can safely handle certain highest average plate currents continuously. This is based on the permissible plate dissipation of the rectifier tube. Average plate currents are determined by applied voltages to the tube electrodes and by the load external to the tube.

5.5.11 Tolerances

Tube manufacturers indicate for their tubes various ratings and characteristic curves showing current-voltage variations under average conditions. Because of complicated conditions arising from chemical and physical phenomena and the relatively small area involved in tube design and construction, variations from manufacturers' data frequently occur. If the variations are too large, the tubes are rejected. The permitted range of variations, or the tolerance, of the tubes may easily be as high as 30 percent from average on either side. Furthermore, the characteristics of a tube change with use. Within certain limits, it is still usable. The tolerance allowable for any particular tube depends, of course, on its function in a particular circuit. In some cases, wide tolerance are acceptable. In other cases, tubes must be constructed to very exacting specifications.

5.6 CHARACTERISTICS

Characteristics are used to identify the distinguishing electrical features and values of an electron tube. These values are usually shown in curve form called characteristic curves and may be used for the determination of tube performance and the calculation of additional tube factors.

The electron tube characteristics are obtained from electrical characteristics of a tube in different circuits under certain definite voltage conditions. The conditions of measurement, static or dynamic must be taken into account in order to further describe the tube characteristics. The static characteristics are the values obtained when different d-c potentials are applied to the tube electrodes, while the

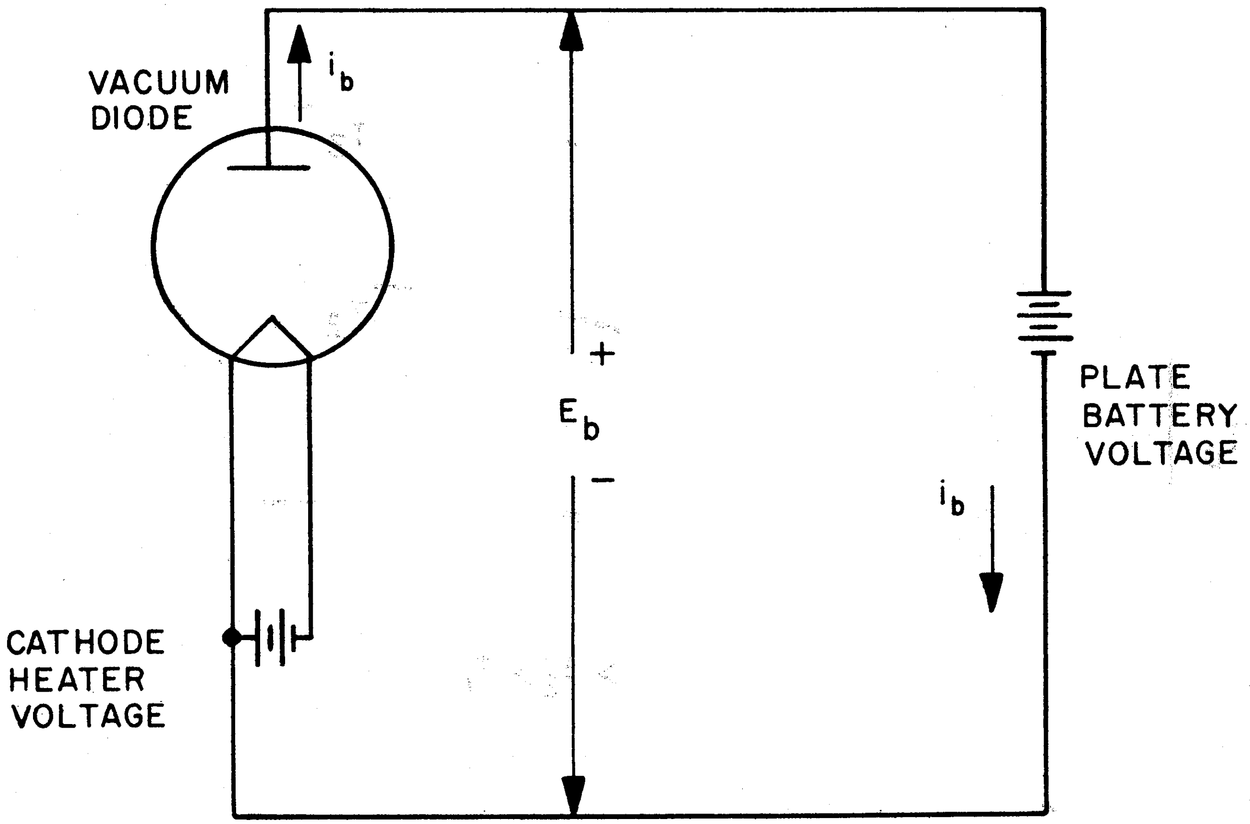

VACUUM
DIODE
i_b
+
E_b
−
PLATE
BATTERY
VOLTAGE
i_b
CATHODE
HEATER
VOLTAGE

Figure 3-90

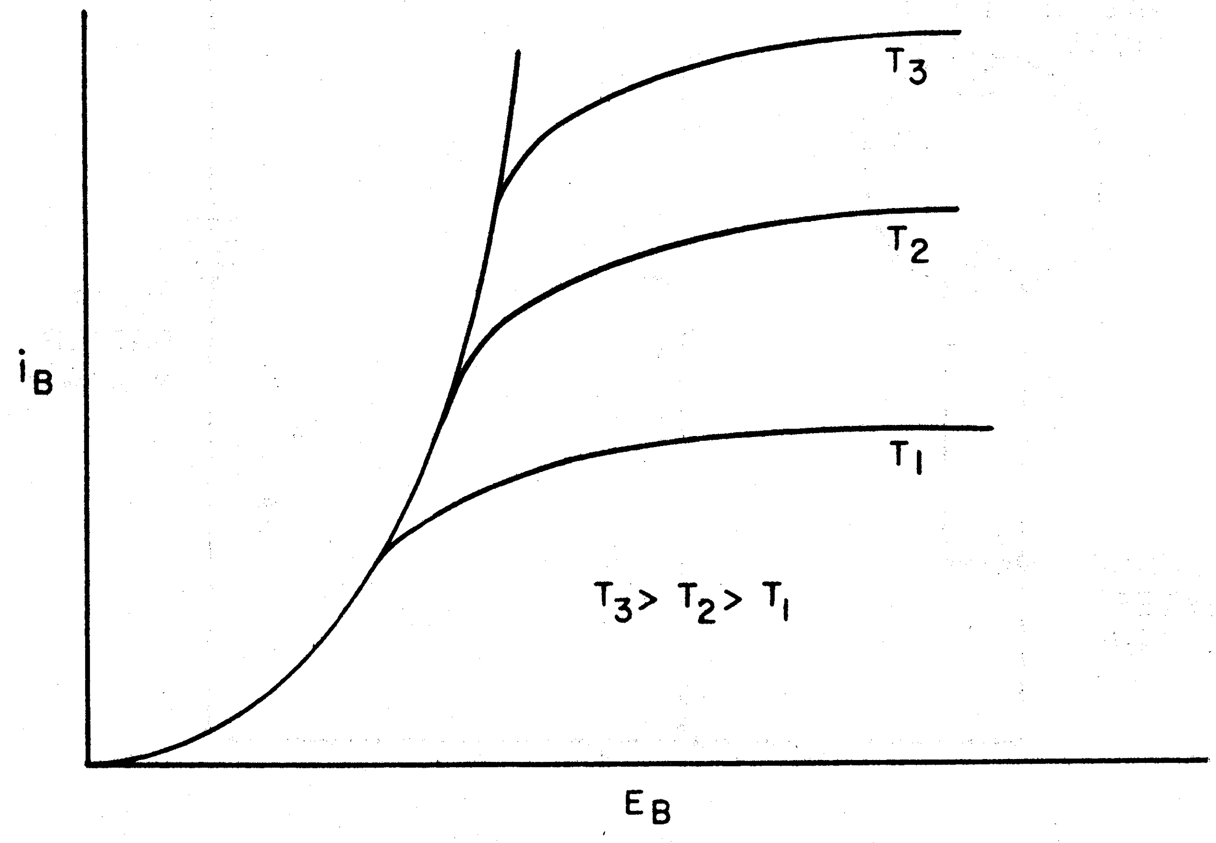


Figure 3-91

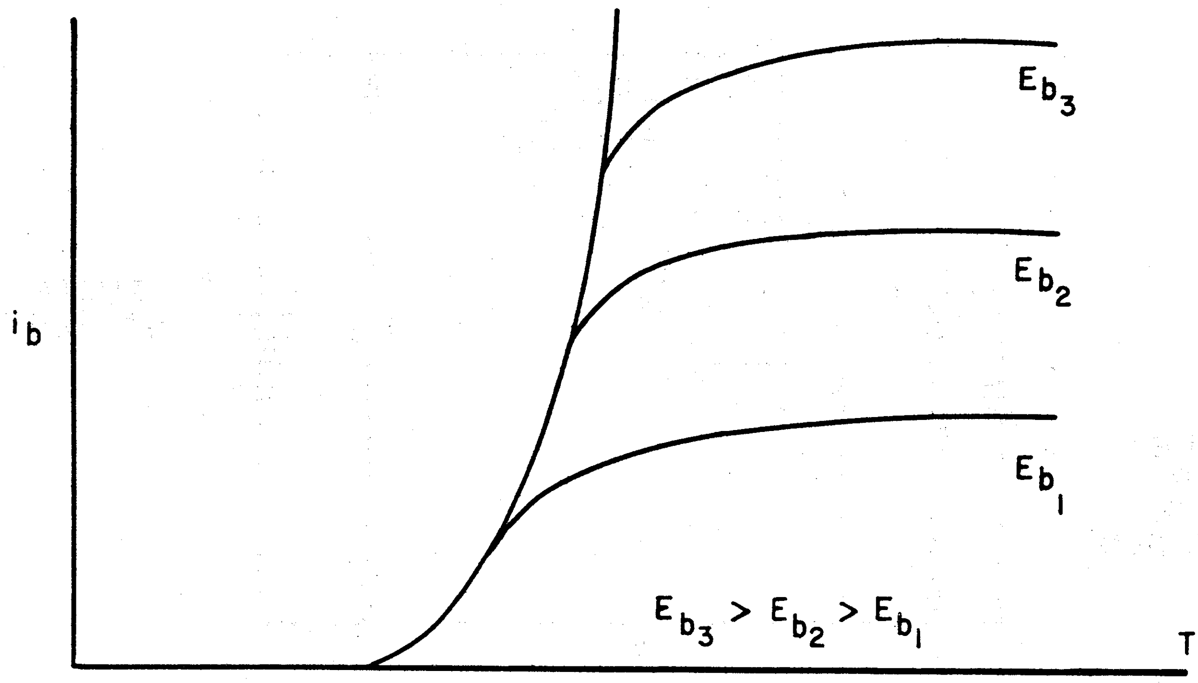


Figure 3-92

dynamic characteristics are the values obtained whan an ac voltage is applied to the control grid under different conditions of dc potential on the electrodes. The performance capabilities of a tube under actual working conditions are demonstrated by the dynamic characteristics.

The characteristic curves indicate the relationship between voltage on one electrode and that on another while other factors are kept constant or between a current and a voltage while still other factors are constant. This section will deal with characteristic curves for the tubes discussed in section 5-3.

5.6.1 Vacuum Tubes

5.6.1.1 Diode

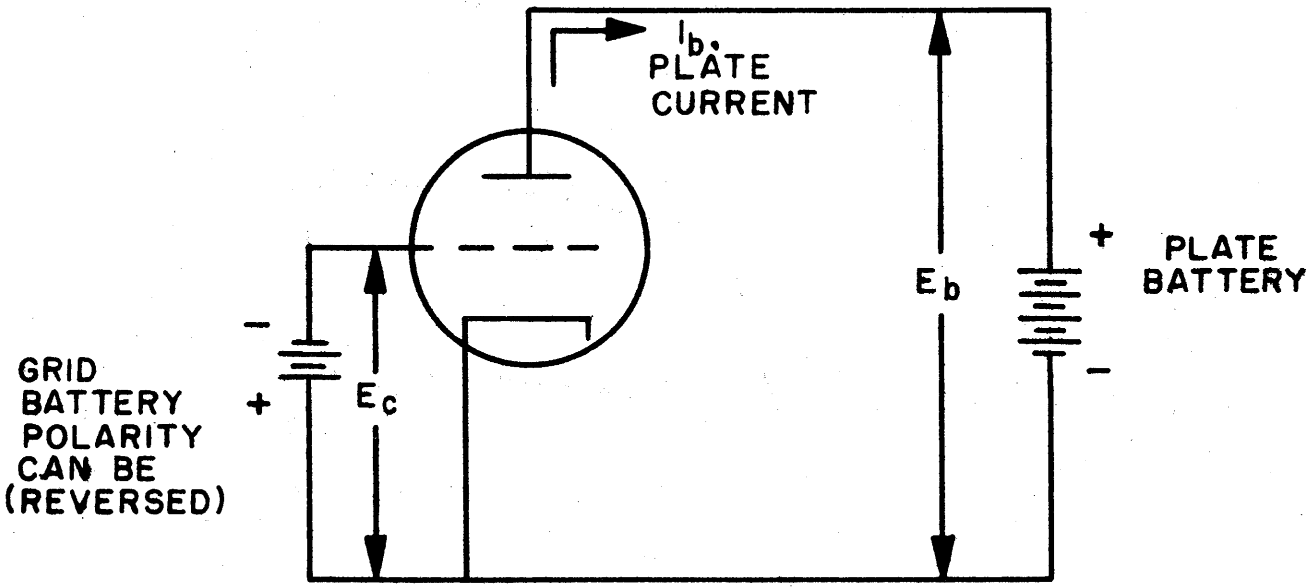
Figure 3-90, shows a typical circuit which might be used to obtain diode characteristic curves. It is assumed that the plate battery voltage and the cathode heater voltage may be varied. The current flow through the tube and in the external circuit is indicated by i_b, the voltage drop across the tube by e_b. Note that the diode shown here has a directly heated cathode.

The characteristic curves for this diode are ideally indicated in Figure 3-91 and Figure 3-92. For given settings of cathode heater voltage, the temperature of the cathode heater is indicated as T_1, T_2, or T_3. In Figure 3-91 plate current i_b for a given value of T increases as the plate voltage e_b increases until the value of i_b is reached where substantially all the electrons emitted by the cathode at that particular temperature are drawn to the plate. Further increases of e_b cannot draw more electrons. If, however, the cathode heater voltage is raised, more electrons can be emitted for the higher temperature, and the maximum plate current is increased. For any value of T, the current has a saturation value indicated by the flat portion of the curve. For currents below saturation, the flow of electrons to the plate depends on both the plate voltage and the space charge formed near the cathode by electrons emitted but not drawn to the plate. At saturation, space charge effects are substantially overcome.

Figure 3-92 shows the variation of plate current with temperature of the cathode for different fixed values of the plate voltage e_b. For any particular value of e_b, more and more electrons are drawn to the plate as the temperature of the cathode is increased until the point is reached where that particular value of e_b cannot drawn more electrons because of the space charge effect. This space charge effect repels electrons being emitted cathode, thus counterbalancing the plate voltage effect. If, however, plate voltage e_b is increased, a greater maximum plate current is possible.

In practice, the curves are not as idealized as indicated. Instead, they separate slightly in the lower portions of the rise of the curves.

The diode characteristics for indirectly heated cathodes are similar in shape to the curves in Figures 3-91 and 3-92



I_b, PLATE CURRENT
+
PLATE BATTERY
–
E_b
–
+
GRID BATTERY POLARITY CAN BE (REVERSED)
E_c

Figure 3-93

Oxide-coated cathodes emit electrons so freely that the limitation on plate current is not set by the temperature of the heating element. Space charge is the only current limiting factor in this tube. The temperature T is usually referred to the absolute or Kelvin scale.

5.6.1.2 Triode

Plate current in a triode depends on the plate voltage, the grid voltage, and the cathode heating voltage. Assuming that the heating voltage is more than sufficient to emit all the electrons which can possibly be drawn to the plate, the plate current will then depend only on the plate and the grid voltages. The three factors under consideration here are i_b, the plate current, e_b, the plate voltage, and e_c, the grid voltage. Three families of curves are used, each one containing a different two of the three factors as variables. These families of curves are called static characteristics. When any two factors are used on a family of static characteristics, the third factor is constant for any particular curve.

Figure 3-93 shows a circuit which might be used to obtain triode static characteristics. It is assumed that the plate voltage and the grid voltage can be varied as necessary and that each of these voltages, as well as plate current, can be measured by appropriate meters which are not indicated in the diagram. A battery for heater voltage is not shown in this diagram.

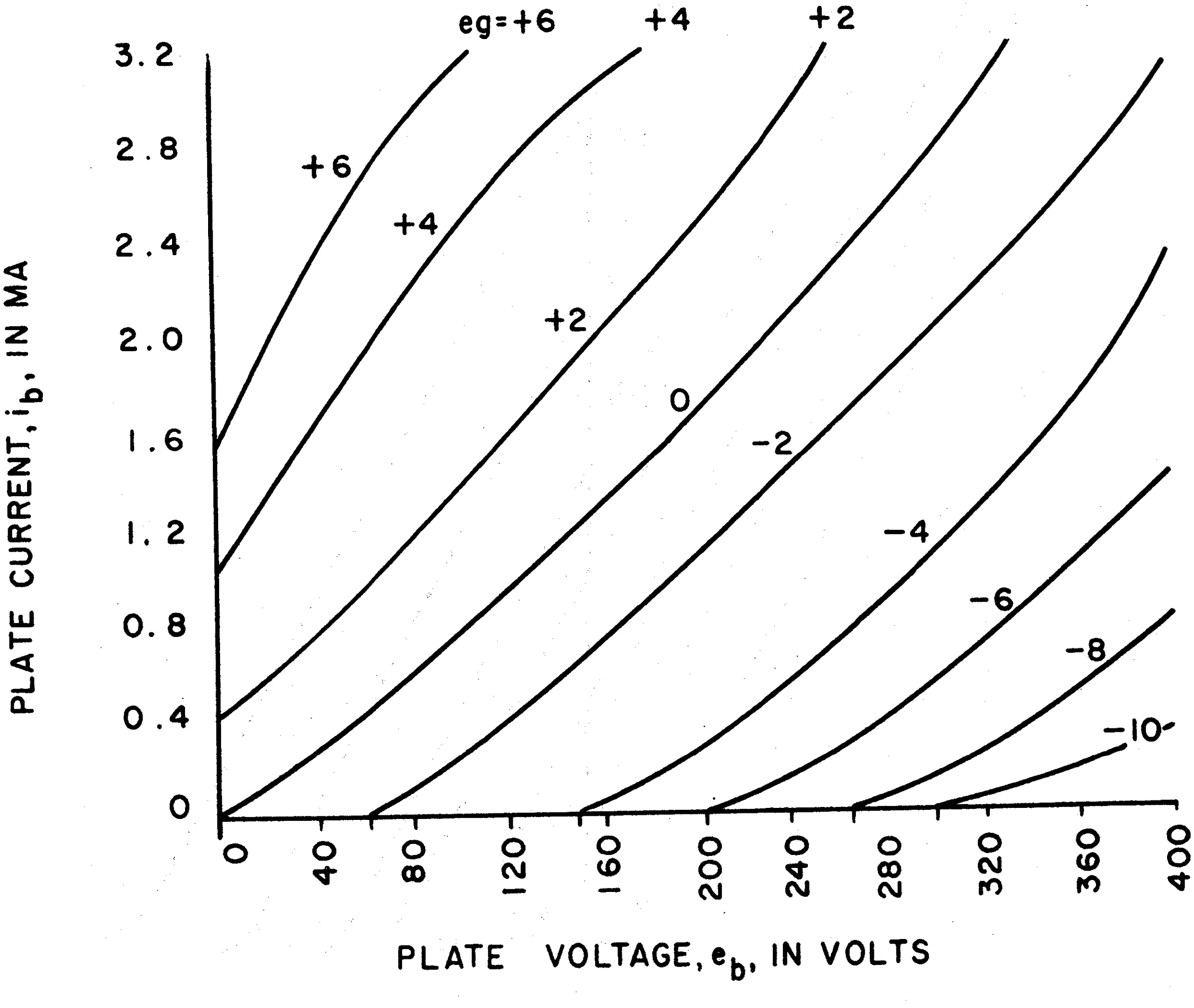
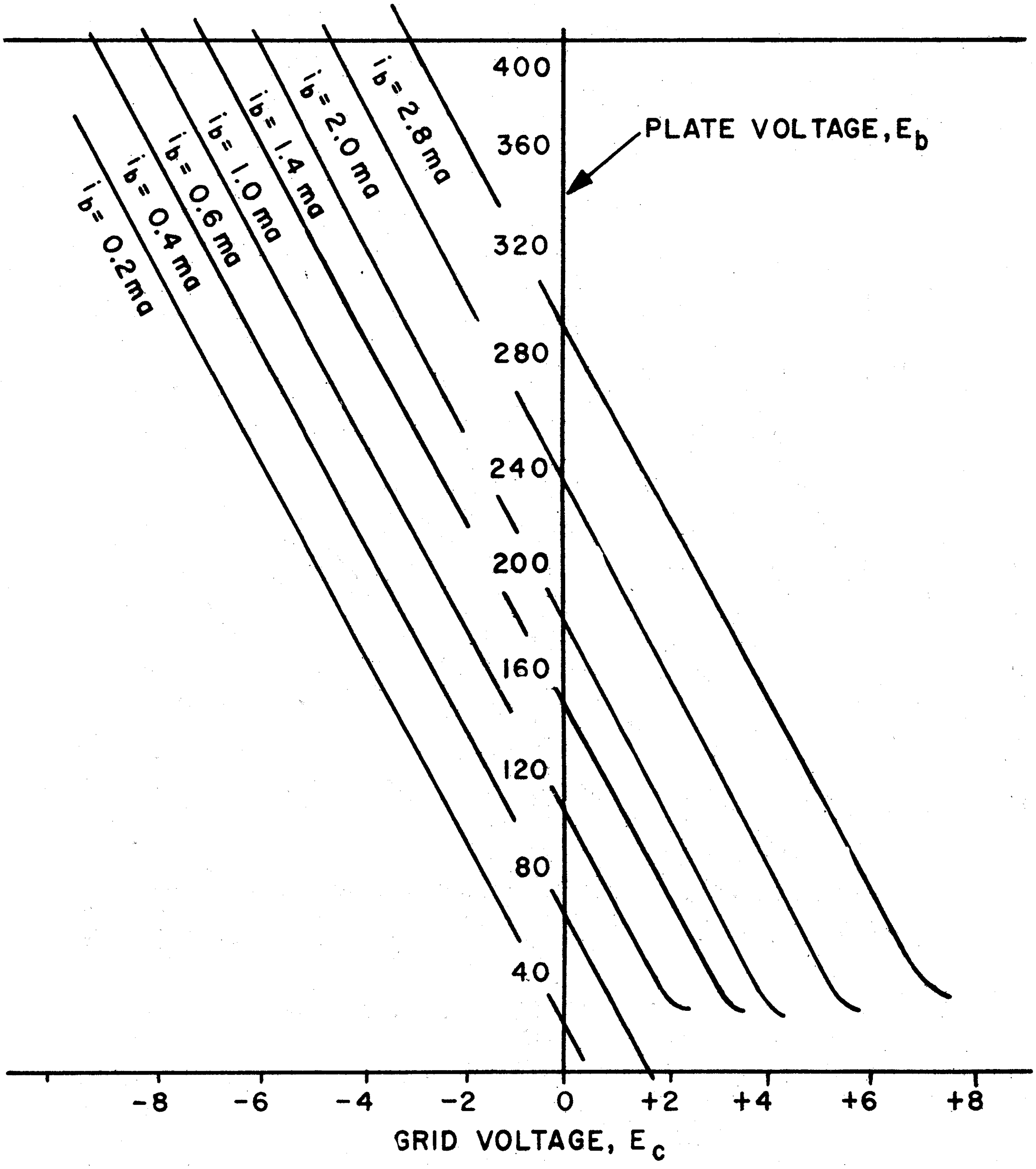

eg=+6
+4
+2
+6
+4
+2
0
-2
-4
-6
-8
-10
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
PLATE CURRENT, i_b, IN MA
0
40
80
120
160
200
240
280
320
360
400
PLATE VOLTAGE, e_b, IN VOLTS

Figure 3-94



i_b = 2.8 ma
i_b = 2.0 ma
i_b = 1.4 ma
i_b = 1.0 ma
i_b = 0.6 ma
i_b = 0.4 ma
i_b = 0.2 ma
PLATE VOLTAGE, E_b
400
360
320
280
240
200
160
120
80
40
-8
-6
-4
-2
0
+2
+4
+6
+8
GRID VOLTAGE, E_c

Figure 3-95

Figure 3-94 shows the plate-current versus plate-voltage characteristics. These are called the plate characteristics. In this family of curves, the grid voltage is constant for any one curve. In the curve where the grid potential has a constant voltage of 0 volts, the triode acts effectively as a diode. As plate voltage increases, plate current increases, assuming saturation current is not reached. If, however, a fixed negative potential is placed on the grid, a higher plate voltage is needed to draw electrons from the cathode. If a positive potential is one the grid, a smaller plate potential is needed to draw electrons from the cathode. Because of the closeness of the grid to the cathode, a small potential change on the grid has a much larger effect on plate current than the same potential change on the plate. Thus, referring to Figure 3-94, if the grid potential is -2 volts, a plate voltage of 240 volts yields a plate current of about 1.42 milliamperes. If the grid potential should change to -4 volts, a decrease of only 2 volts, then a plate voltage of about 340 volts is required to yield the same plate current of about 1.42 milliamperes. Thus, a change here of two volts on the control grid is counterbalanced by a change of 100 volts on the plate.

Figure 3-95 shows a family of plate-voltage versus grid-voltage characteristics for a series of different constant plate currents. These curves are often called constant current characteristics. They show, for any particular constant current, the change in plate voltage required to meet a change in grid

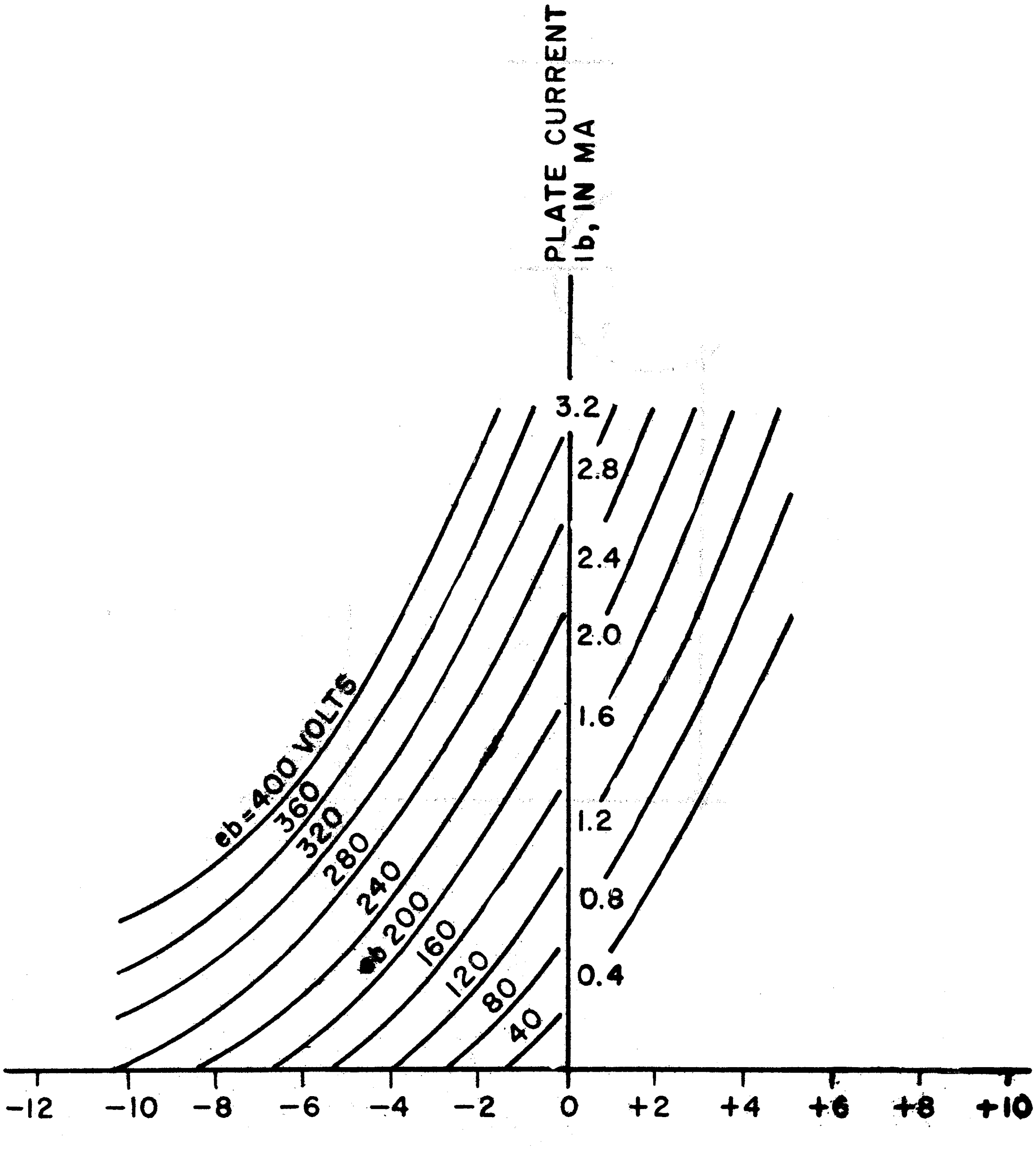


PLATE CURRENT
Ib, IN MA
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
eb=400 VOLTS
360
320
280
240
eb 200
160
120
80
40
-12
-10
-8
-6
-4
-2
0
+2
+4
+6
+8
+10
GRID VOLTAGE, e, c,

Figure 3-96

voltage in order to keep the current constant. Note that these curves have negative slopes indicating that changes in plate voltages have **opposite** effects from changes in grid voltage.

Figure 3-96 shows a family of plate-current versus grid voltage characteristics for different fixed values of plate voltages. These curves are often called transfer characteristics.

Any one of all the characteristics shown in **Figures** 3-94, 3-95 and 3-96 for a particular triode can be obtained from any other. Each shows the same information in a different form.

5. 6.1.2.1 Vacuum Tube Coefficients

The characteristic curves can readily be used to analyze the behavior of a tube. An alternative method using tube coefficients is often useful. This method with coefficients is, however, approximate but permits application in circuits analysis in those regions where the curves are linear.

Plate resistance, r_p, is defined as the ratio of plate voltage change, Δe_b, to the plate current change, Δi_b, when grid voltage, e_c, is kept constant. Thus:

$$r_p = \frac{\Delta e_b}{\Delta i_b} \qquad e_c = \text{constant.}$$

It is the reciprocal of the slope along the straight portions of the plate characteristics in **Figure** 3-94. It is assumed these straight portion slopes are all the same. Actually, they are not. The plate resistance r_p is also known as the dynamic, incremental, or variational plate resistance. It should not be confused with $\frac{e_b}{i_b}$, called the static or d-c plate resistance.

The transconductance, g_m, is the ratio of the change in plate current, Δi_b, to the change in grid voltage, Δe_c, assuming plate voltage, e_b, is kept constant.

$$g_m = \frac{\Delta i_b}{\Delta e_c} \qquad e_b = \text{constant.}$$

Transconductance is the slope of the linear portions of the transfer characteristics **in Figure** 3-96.

The amplification factor, μ, is the ratio of change in plate voltage Δe_b to the change in grid voltage Δe_c, assuming plate current i_b is constant.

$$\mu = -\frac{\Delta e_b}{\Delta e_c} \qquad i_b = \text{constant}$$

The negative sign is used because these voltage changes have opposite effects, one being positive when the other is negative. The amplification μ is a positive number; referring **to Figure** 3-94, the constant current characteristics, μ, is the negative of the slope of the curves along their linear portions. These slopes are negative, but μ, again, is positive.

It can be proved mathematically that

$$u = g_m r_p$$

It is important to note that these three coefficients are constants, and therefore useful only in the regions where the curves **in Figures** 3-94, 3-95, **and** 3-96 **are straight**, parallel, and equidistant for equal changes in e_b, e_c, or i_b whichever applies for the particular static characteristics.

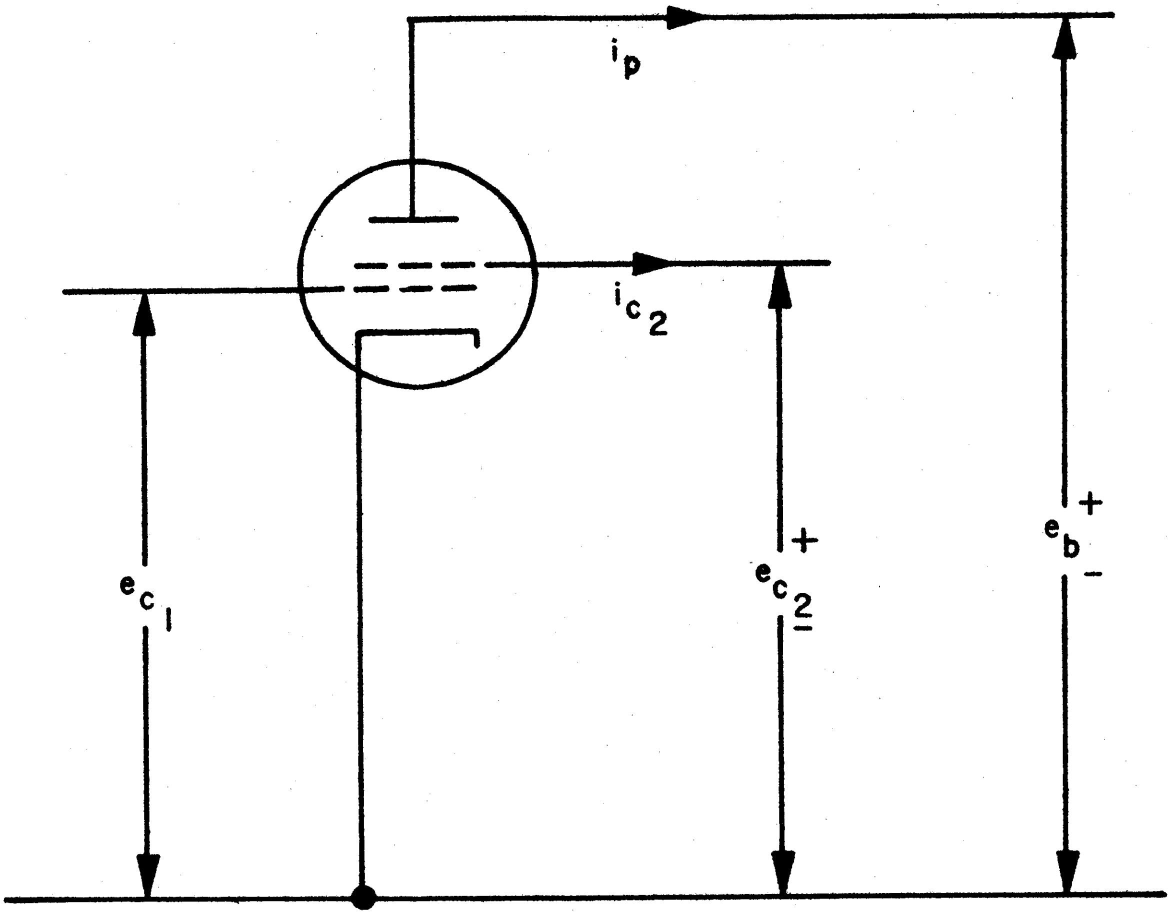


Figure 3-97

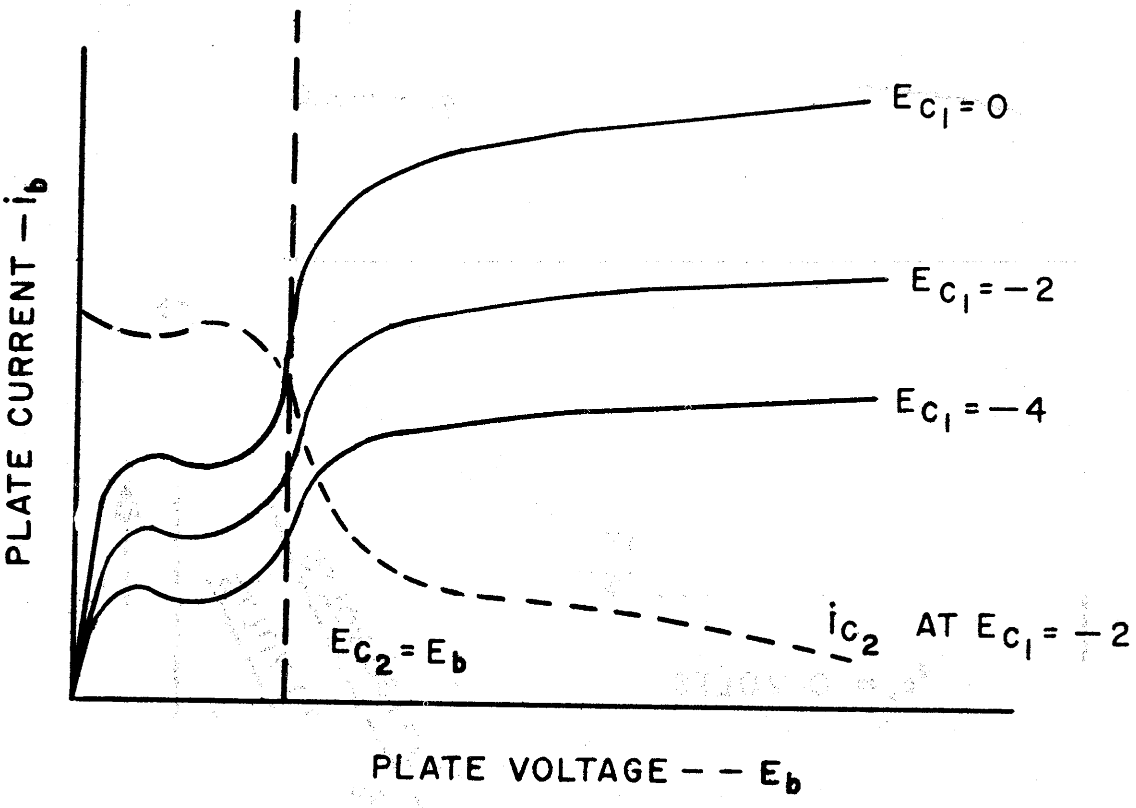


PLATE CURRENT — i_b
$E_{c_1} = 0$
$E_{c_1} = -2$
$E_{c_1} = -4$
i_{c_2} AT $E_{c_1} = -2$
$E_{c_2} = E_b$
PLATE VOLTAGE -- E_b

Figure 3-98

5.6.1.3 Tetrode

Figure 3-97 shows a partial tetrode circuit. The grid nearest the cathode is referred to as grid 1. It is the control grid. Usually, it is the only grid allowed to vary in voltage to control current. The voltage on it with respect to the cathode is called e_c. Grid 2 is the screen grid. It usually has a fixed voltage in any particular circuit. This voltage is e_{c_2}. The effect of the screen grid is to shield the control grid from the plate and thereby reduce the capacitance between them.

Figure 3-98 shows typical plate characteristics for a tetrode. Note that the screen grid voltage e_{c_2} is the same for all three curves shown. Each of the three curves shows the variation of plate current with plate voltage for different values of control grid voltage e_{c_1}. The particular value of e_{c_2}, screen grid voltage, is indicated on Figure 3-98, by the vertical dotted line. The characteristics of plate current versus plate voltage are the ones most commonly shown for tetrodes.

To explain the reasons for the shape of the curves, consider the one for a fixed control grid voltage $e_{c_1} = 2$ volts. As plate voltage e_b increases from zero, plate current i_b increases. For plate voltages up to point p (see Figure 3-98), the primary electrons striking the plate have such low velocities that few secondary electrons are emitted from the plate. As plate voltage increases beyond point p, more and more secondary electrons are emitted from the plate. In fact, more are emitted for values of e_b from p to q than strike the plate. This causes a reverse conduction current. These secondary electrons are attracted by the higher screen grid voltage e_{c_2} and become part of the screen grid current i_{c_2} (To the left of the vertical dotted line, e_{c_2} is greater than e_b.) Moving beyond point q, the plate voltage e_b rises and reduces the velocity of secondary electrons going toward the screen grid. The net effect is a rising plate current i_b until the point where $e_b - e_{c_2}$. Here the plate and screen voltages are equal and relatively few secondary electrons **from** the plate reach the screen grid. Beyond the point where $e_b = e_{c_2}$ the plate voltage has little effect on increasing the plate current because the screen grid effectively shields the plate from the cathode. The current, therefore, levels off.

The screen grid, being positive with respect to the cathode, draws current also. This is indicated by the dotted curve labeled i_{c_2} in **Figure** 3-98. The sum of plate and screen grid currents is practically constant.

Where tetrodes are used, operation is generally restricted to the flat portions of the characteristics where plate voltage is greater than screen voltage.

5.6.1.4 Pentode

Pentodes are widely used for purposes of amplification and have generally replaced tetrodes. When a pentode is used as an amplifier, grid 1 is used as a control grid to control plate current, grid 2 is held at a fixed positive voltage to screen the control grid from the plate, and grid 3 is used to

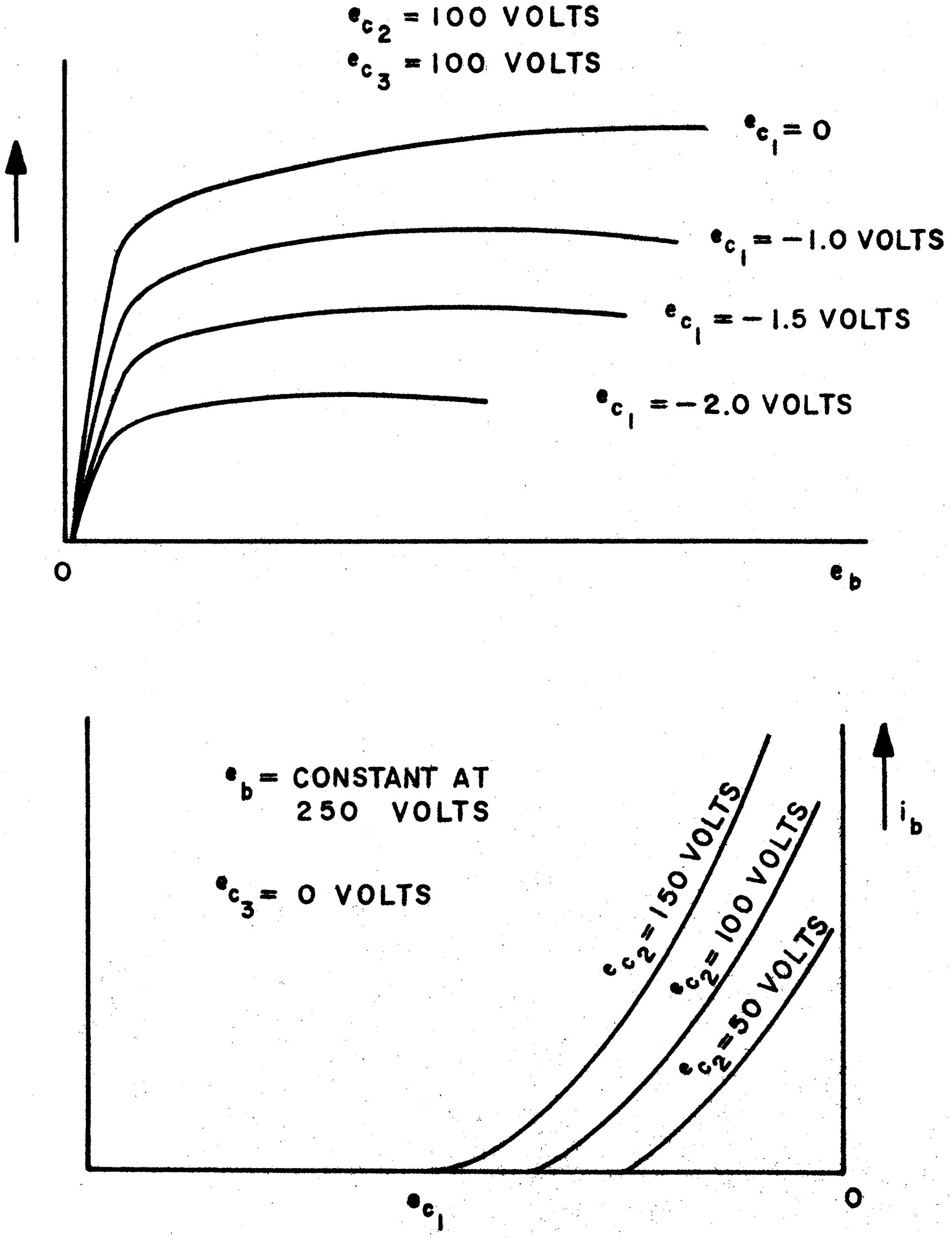

e_{c_2} = 100 VOLTS
e_{c_3} = 100 VOLTS
e_{c_1} = 0
e_{c_1} = −1.0 VOLTS
e_{c_1} = − 1.5 VOLTS
e_{c_1} = −2.0 VOLTS
0
e_b
e_b = CONSTANT AT 250 VOLTS
e_{c_3} = 0 VOLTS
e_{c_2} = 150 VOLTS
e_{c_2} = 100 VOLTS
e_{c_2} = 50 VOLTS
i_b
e_{c_1}
0

Figure 3-99

suppress secondary emission from the plate. This suppressor grid, usually at a fixed voltage of zero with respect to the cathode, provides for the elmination of the dip in the characteristics for a tetrode shown in **Figure** 3-98. Pentodes have very high amplification factors. Figure 3-99 **shows** typical plate and transfer characteristics for a pentode.

5.6.1.5 Beam Power Tube

Beam power tubes were developed after tetrode and pentode tubes. Beam power tubes have two grids, a control grid and a screen grid. The design is such, however, that the electrons move from cathode to plate in sheets or flat beams. The space charge formed between screen grid and plate is so negative as to **suppress** the effects of secondary emission from the plate. The effect is virtually the same as if a suppressor grid were present.

Plate characteristics of the beam power tube are like those for a pentode except that the knees of the curves are sharper and the horizontal portions of the curves cover a wider range of plate voltages. This permits greater power output and efficiency.

Also, in a beam power tube, the grids of the screen are aligned with those of the control grid to permit the forming of electron sheets. This makes for a smaller screen grid current in proportion to the plate current than **occurs in a** pentode.

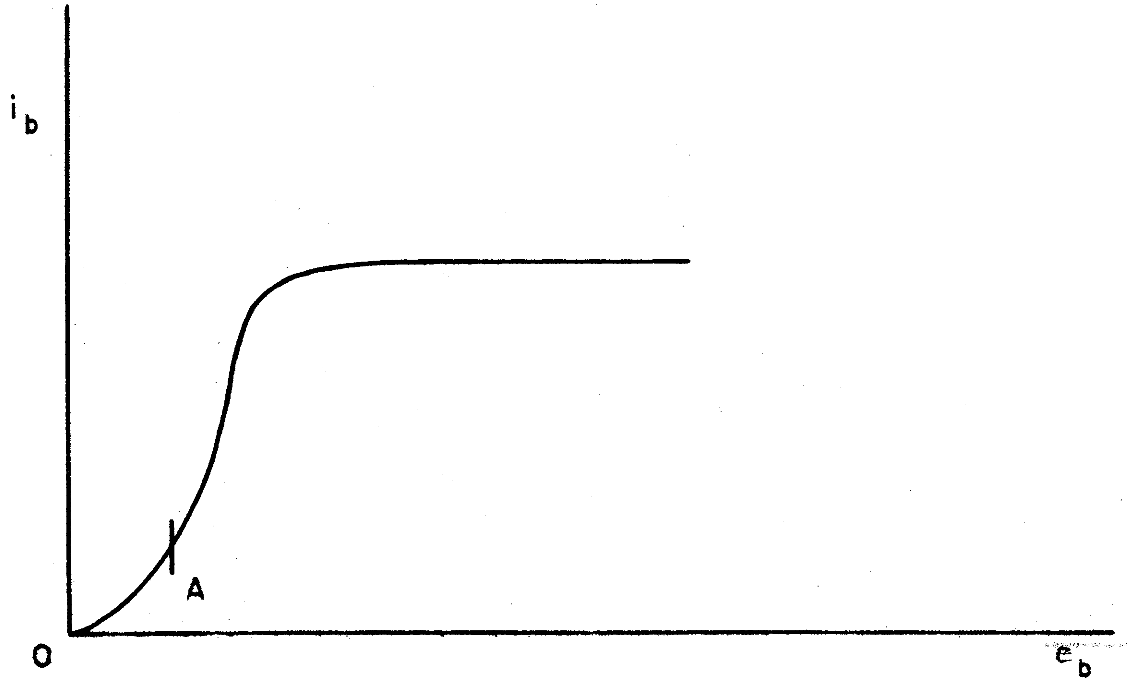


Figure 3-100

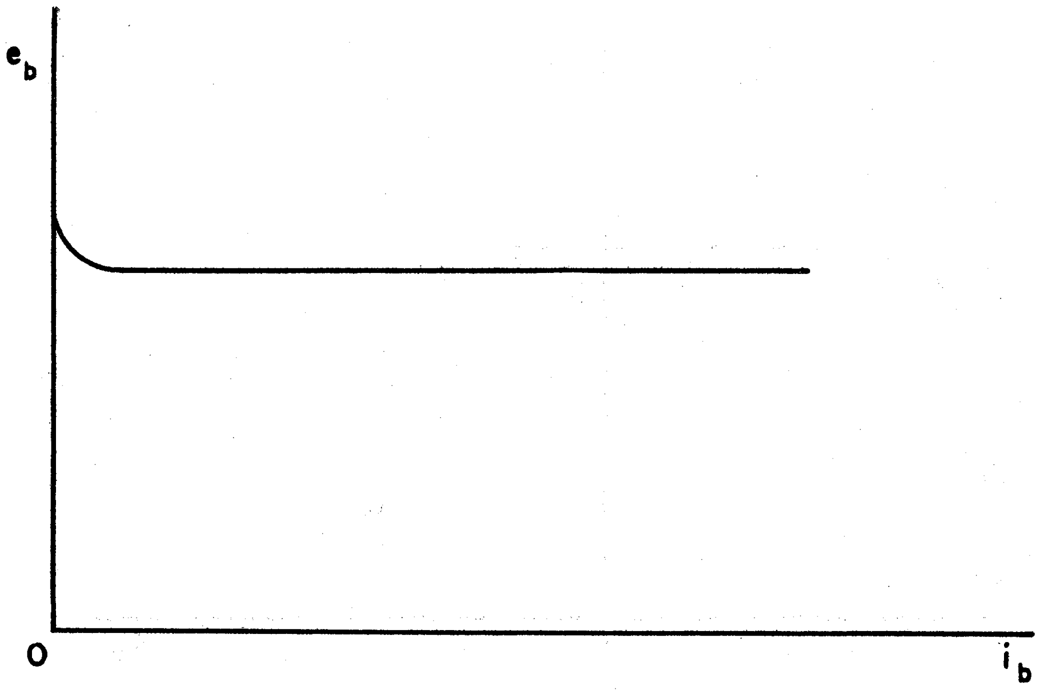


Figure 3-101

5.6.1.6 Multigrid Tubes

So many electrode voltages are involved in multigrid tubes that static characteristics are not generally given. Instead the best operating conditions and performances of these tubes are individually given in the specific applications where they are used.

5.6.1.7 Multi-Unit Tubes

In multi-unit tubes, each individual tube section has its own static characteristics as previously discussed.

5.6.2 Gas Tubes, Thermionic Cathodes

5.6.2.1. Gas Diodes

When the gas pressure in a gas diode with a thermionic cathode is low, the characteristic curve is as shown in Figure 3-100. From O to A, the effect of the gas is small. At A, however, the ionizing potential is reached, and the current increases very rapidly to its maximum value. Thereafter, increases in voltage have no effect on the current.

If the pressure of gas in the diode is high enough, initial increases in voltage from zero have negligible effect on the current. However, for these very small current values, as voltage is increased, the ionizing potential is reached and the tube breaks down, the current rising at once to its maximum value. Figure 3-101, shows the variation of current with voltage in a gas diode, thermionic cathode. Note that i_b and e_b are reversed on the axes. As e_b increases from zero, i_b does not remain zero, but it is so small as not to be perceptible on

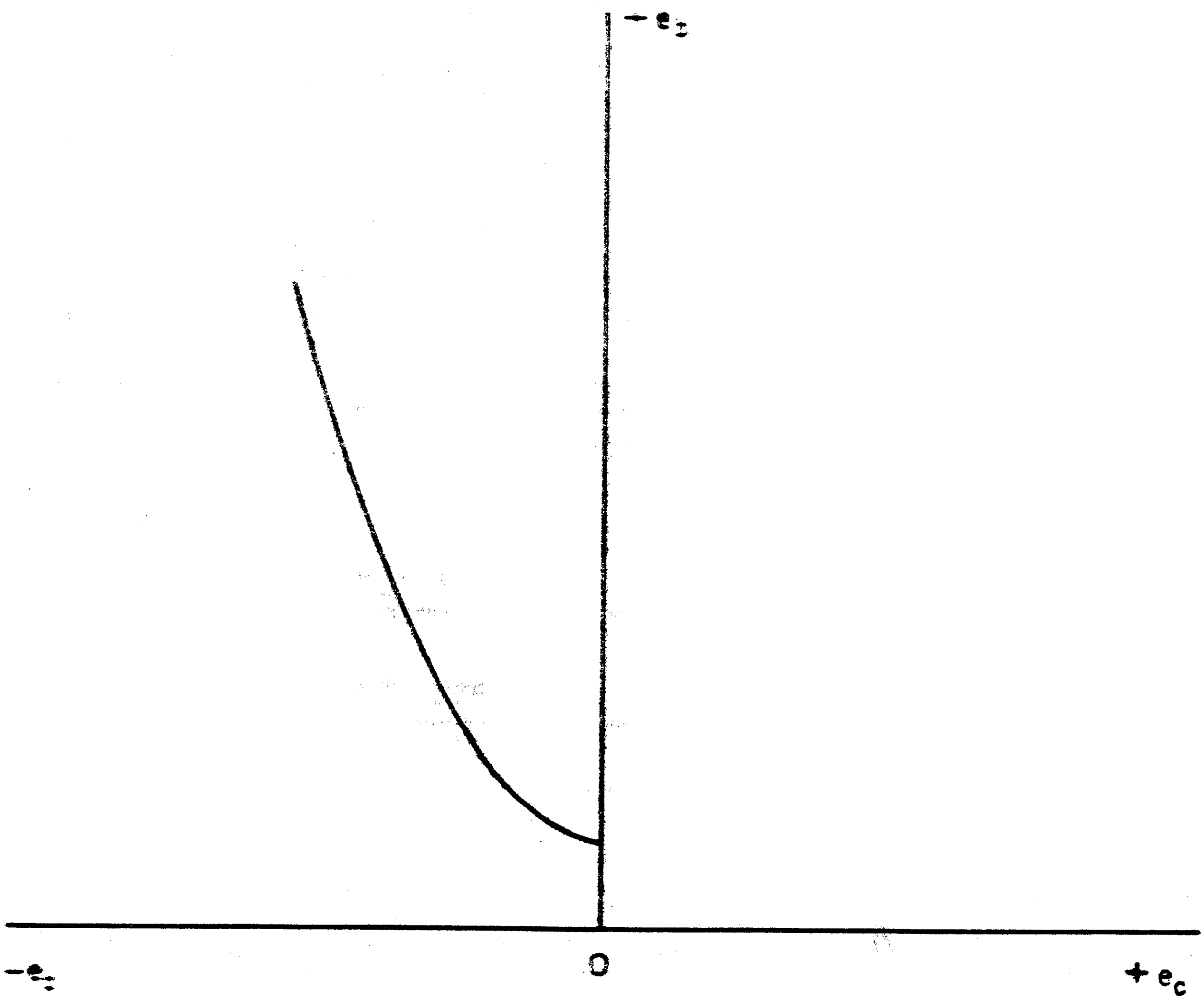


Figure 3-102

the graph. The current is limited by the external circuit resistance assuming cathode emission is sufficiently high, as it is with an oxide-coated cathode. The tube glows when the tube breaks down. The value of e_b after firing is 10 to 20 volts.

In a gas-filled thermionic tube, it is essential to allow the cathode to heat up and become fully emissive before voltage is applied to the tube. Otherwise, if the cathode cannot supply the emissive current demanded by the voltage, the cathode can be ruined by bombardment of positive ions repelled by the positive plate. Such tubes require a time delay between the application of cathode heating voltage and plate voltage.

5.6.2.2 Thyratron

The grid of a thyratron controls the firing point. Grid current of the order of milliammeters is possible because the grid collects electrons or positive ions depending on its potential. To keep grid current low, a high impedance is put in the grid circuit. Once the tube fires, the grid **loses** its effects on controlling current because of the presence of a shield of positive ions about the cathode. The important characteristic in a thyratron is called a starting or control characteristic. This is illustrated in **Figure** 3-102.

A starting characteristic, which depends on the type tube and on the temperature of the gas (for a fixed pressure) establishes a boundary between two regions. On the right and

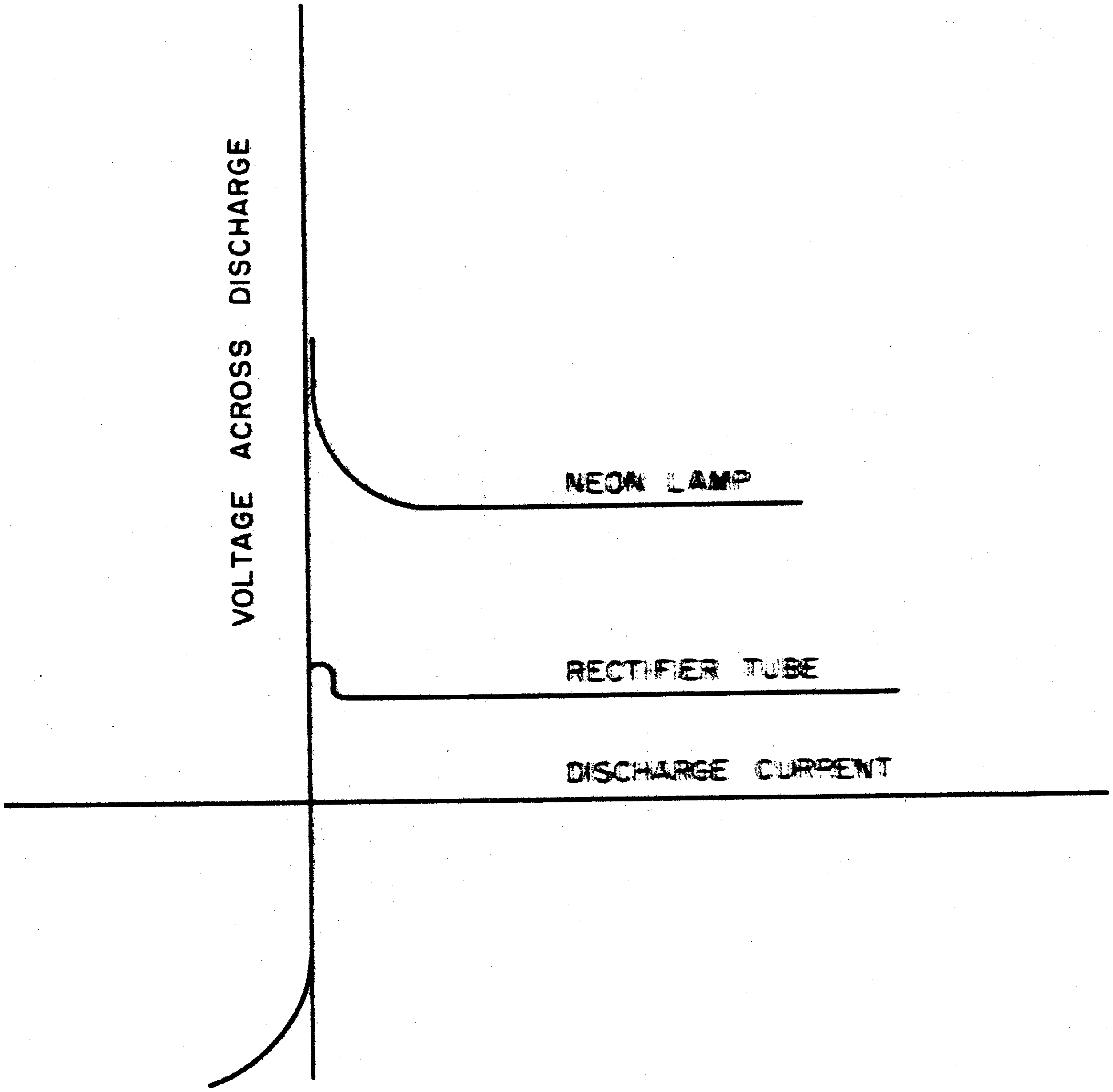

VOLTAGE ACROSS DISCHARGE
NEON LAMP
RECTIFIER TUBE
DISCHARGE CURRENT

Figure 3-103

above the characteristic, the tube conducts. On the left and below the characteristic, it will not conduct. If the tube is not conducting to begin with and the plate voltage is at some fixed value, then approaching the curve from the left along a horizontal line through fixed plate voltage will result in the tube's firing as soon as the curve is reached; i.e., as soon as the proper value of grid voltage is reached. Or, if the grid voltage is fixed, the tube fires as soon as the curve is reached when approached along a vertical line from below. It is important to note that, once the tube is conducting, it cannot be extinguished by approaching the curve from above or from the right.

Thyratrons are classified as negative-control or positive control tubes. Negative-control thyratrons have negative starting grid voltages over substantially the entire range of operations. Positive-control thyratrons have positive starting grid voltages over substantially the entire range of operations.

5.6.3 Gas Tubes, Cold Cathodes

Figure 3-103 shows volt-ampere characteristics for two different types of cold cathode tubes. Note the essentially constant voltage drop for a wide range of current values in the positive region for currents. The cathode generally glows in such a tube. Where it is used, however, with alternating voltage applied to its electrodes, both electrodes glow.

For use as a rectifier, the amplitude of the alternating voltage is higher than the positive portion of the rectifier tube curve (see Figure 3-103) but lower than the value to the negative knee of this curve. In such rectifier tubes, the lack of symmetry in the curves is due to the fact that one electrode is designed to be a better emitter of electrons than the other for the same field strength.

The largest voltage that can be regulated employing the glow discharge is limited to about 150 volts.

5.7 MULTIPLIER PHOTOTUBES

An electron multiplier is a device for taking an initial source of electrons and converting it into a much heavier flow of electrons through a series of successive stages. Each stage takes the number of electrons impinging on it and approximately multiplies this number by some multiplier greater than one. A phototube is a device for converting light energy into electron flow. Light energy impinging on a specially prepared metal surface causes electrons to be emitted by that surface.

In a multiplier phototube, the light impinging on the specially prepared metal surface releases relatively few electrons. By multiplier action, however, these few electrons are increased by multiplication in each of several stages as the electron stream moves from one stage to the next. In each stage, one electron hitting a surface releases several electrons from it by secondary emission. Each of these several electrons then hitting the surface in the next stage

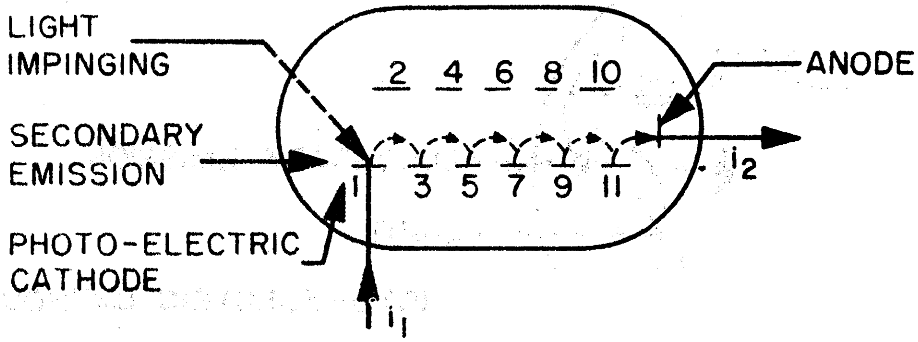

LIGHT IMPINGING
SECONDARY EMISSION
PHOTO-ELECTRIC CATHODE
ANODE
2
4
6
8
10
1
3
5
7
9
11
i_2
i_1

Figure 3-104

again releases several electrons. Ideally, for example, if one electron causes two to be releasedd by any one stage, then in ten stages the number of electrons released for each initial one will be 2^{10} (1024). In practice, some electrons are lost from the desired path and do not release other electrons, or some electrons do not release the assumed number of electrons from a surface.

An early type of photomultiplier tube is shown in **Figure 3.104**. Light impinging on the photoelectric cathode marked 1, releases a few electrons which tend to move up toward the positive electrically charged plate marked 2. Not shown in the diagram is a magnetic field directed perpendicularly to the diagram. The electron stream, constituting a current, is bent at right angles to the magnetic field and to its own stream path. It follows the dotted curve toward the lower plate marked 3. This lower plate, as well as those lower plates marked 5, 7,9, etc, have specially prepared surfaces which release several secondary electrons for each one impinging on it. The increased electron stream from plate 3 tends to move toward positively charged plate 4, which has a greater positive charge obtained from a voltage divider network than does plate 2, but again the magnetic field turns the stream toward plate 5 along the dotted curve. Once more an increased electron stream is obtained to move upward from plate 5. Finally, the anode picks up an electron current stream i_2 much greater (multiplied several times) than the current i_1, initiated by the light impinging on photoelectric

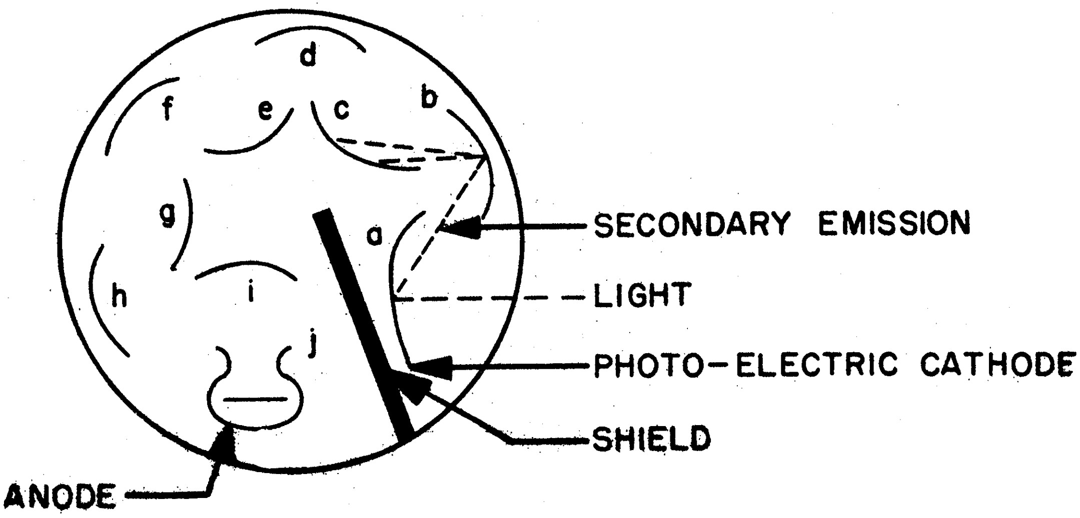

d
f
e
c
b
g
a
h
i
j
SECONDARY EMISSION
LIGHT
PHOTO-ELECTRIC CATHODE
SHIELD
ANODE

Figure 3-105

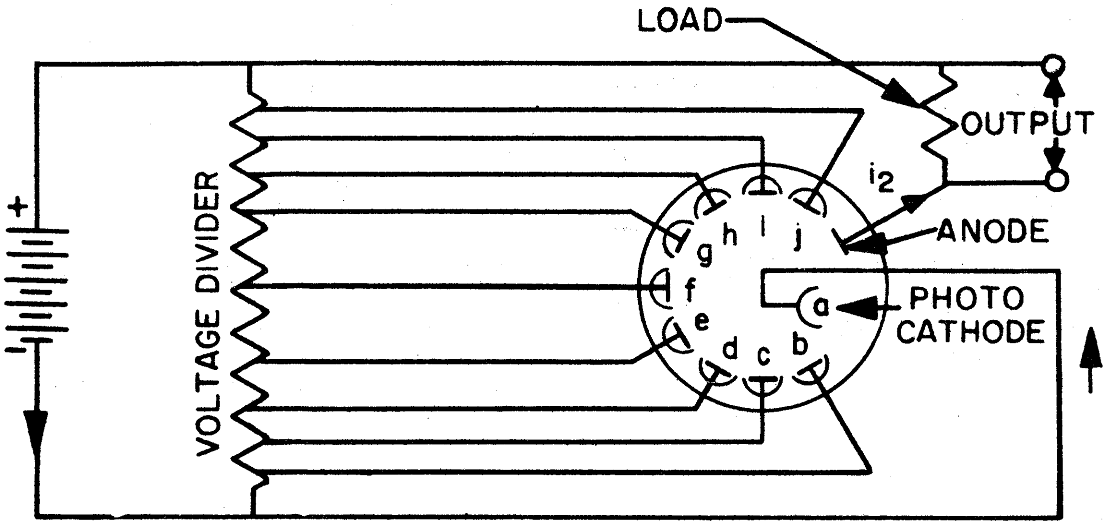

LOAD
OUTPUT
i_2
ANODE
PHOTO CATHODE
i_2
+
-
VOLTAGE DIVIDER
a
b
c
d
e
f
g
h
I
j

Figure 3-106

cathode 1.

A diagram of a common type of electrostatic phototube multiplier is shown in Figure 3-105. The photocathode is marked "a" and serves as the initial source of electrons released by light impinging on it. The secondary emitters are called dynodes and are labelled "b", "c", "d", "e", "f", "g", "h", "i", "j". Each successive dynode is maintained at an increased potential from that of the previous one by means of taps on a resistance voltage divider. Photoelectrons emitted by photocathode "a" are attracted to dynode "b" by the higher potential on it. For each electron hitting a dynode, assume, ideally, that is releases two secondary electrons. Then dynode "b" emits two electrons for each one hitting it. Each emitted electron from dynode "b" is attracted to dynode "c", which has a higher potential than dynode "b". Dynode "c" then emits four electrons. This continues until the electron stream goes from dynode "j" to the anode, which it almost encloses. The dotted lines from "a" to "b" to "c" to "d" indicate how the electron stream increases. One dotted line goes from "a" to "b", two from "b" to "c", and four from "c" to "d".

Figure 3-106, shows a typical schematic representation of a circuit containing an electronic multiplier phototube.

5.8 CATHODE-RAY TUBES

A cathode-ray tube is a special type of vacuum tube in which electrons emitted from a cathode are moved at very high speeds, are formed into a very narrow beam, and then allowed to strike a chemically prepared glass screen which fluoresces,

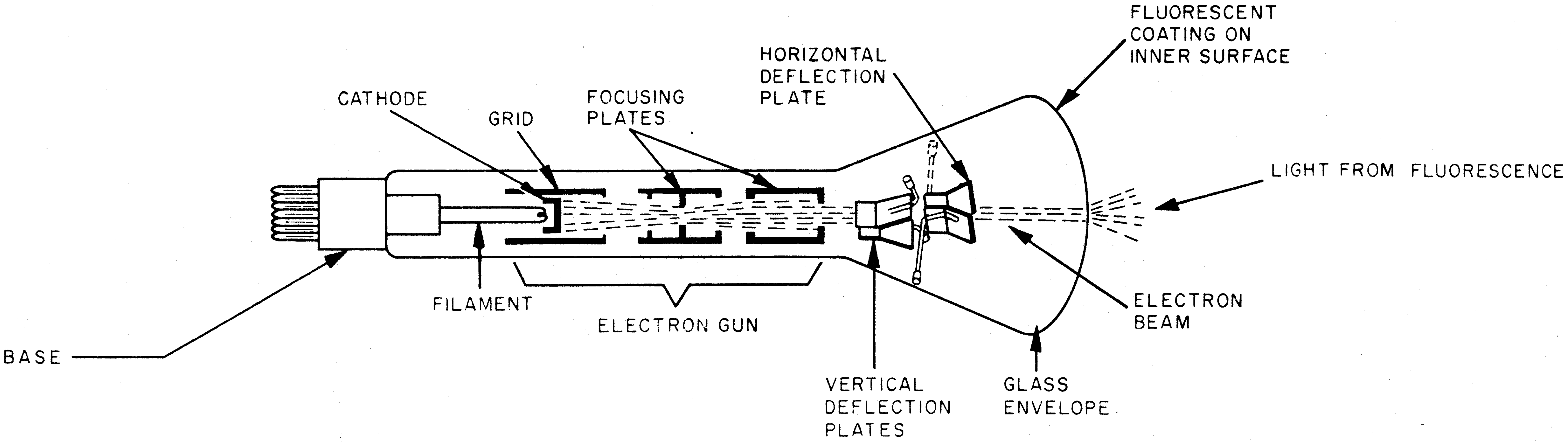

FLUORESCENT COATING ON INNER SURFACE
HORIZONTAL DEFLECTION PLATE
CATHODE
GRID
FOCUSING PLATES
LIGHT FROM FLUORESCENCE
FILAMENT
ELECTRON GUN
ELECTRON BEAM
BASE
VERTICAL DEFLECTION PLATES
GLASS ENVELOPE

Figure 3-107

or glows, at the point where the electron beam strikes. The importance of the cathode-ray tube is that it provide a visual means of examining and measuring current and voltage phenomena in electric circuits. Also, in somecomputers, cathode-ray tubes are used to store binary information with a provision for random access to any bit.

The cathode-ray tube may be considered to have four parts: a heated cathode as a source of electrons; an electron gun consisting of an arrangement of electrodes which serves to accelerate the electrons from the cathode, to focus them into a fine pencil or beam of rays, and to project them down the major axis of the tube; an arrangement of electrodes or magnetic coils located beyond the electron gun and used to deflect the electron beam up or down and right or left; and a glass target or screen, chemically treated, in a plane perpendicular to the major axis of the tube used to provide a visual indication of where the electron beam strikes. Figure 3-107 shows a simplified construction of a typical cathode-ray tube with electrostatic deflection plates. The parts of the cathode-ray tube will be discussed below.

5.8.1 Cathode Emission

Electrons are emitted thermionically from a heated cathode in the cathode ray tube. The beam of electrons, called a cathode ray beam, is relatively wide and electrons move off in several different directions at relatively low velocities. This beam is made pencil narrow by the electron gun which also accelerates the electrons.

5.8.2 Electron Gun

The electron gun consists of a control grid, accelerating anodes, and focusing devices.

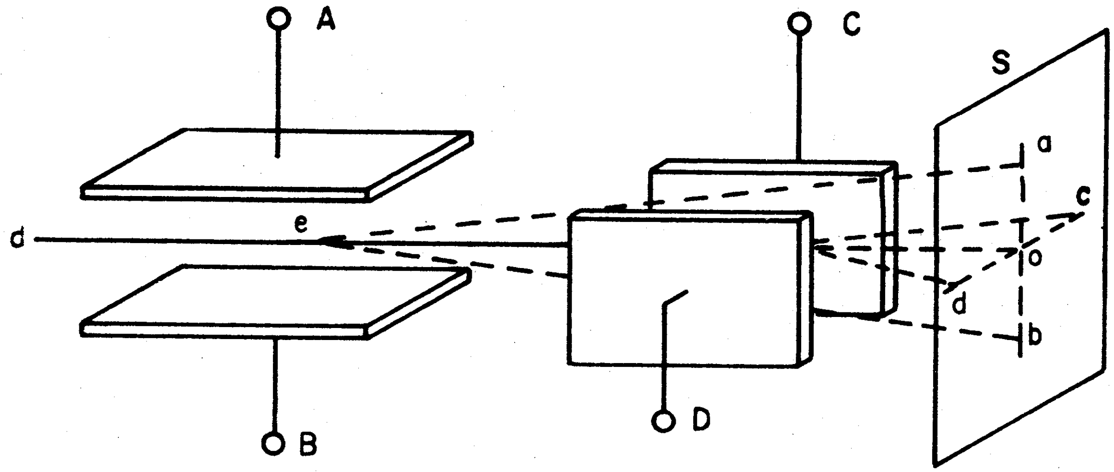
5.8.2.1 Control Grid

After leaving the cathode, the electron beam is narrowed by the extent to which it can pass through the hole of the cylindrical grid which surrounds the cathode. (See **Figure** 3-107) By varying the negative voltage on the control grid, the intensity of the cathode beam is varied.

5.8.2.2 Focusing Anodes

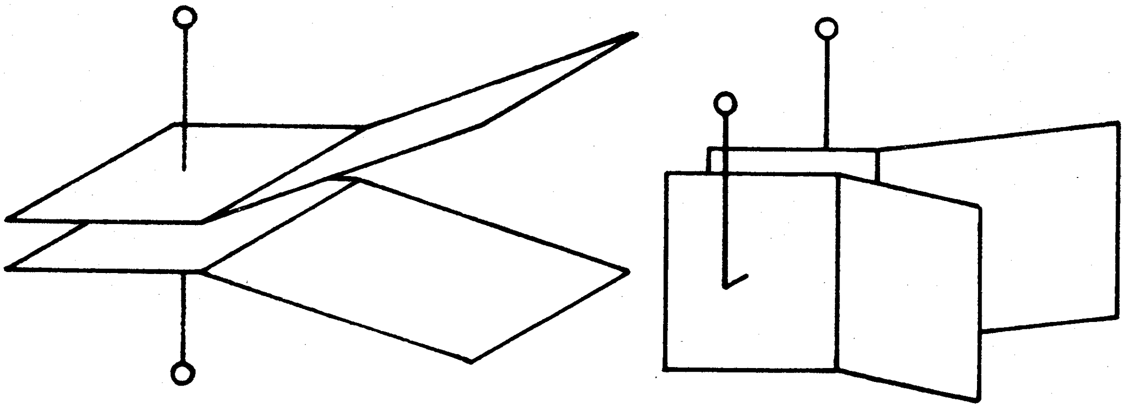
The beams passing beyond the hole in the cylindrically shaped grid are still slow moving and relatively unfocused. The two anodes A_1, and A_2, both cylindrically shaped, serve to focus the beams sharply and to accelerate their motion. Both anodes are positive with respect to the cathode. Anode A_1, the first anode, serves as a focusing anode. Anode A_2, more positive than A_1, serves to accelerate the electrons. The important feature in focusing in the area inside the two anodes is the ratio of positive voltages on these anodes to each other and the resulting shaped electrostatic field established inside the anodes. Leaving the hole of the second anode is a fast moving extremely sharp cathode beam, assuming the proper ratio of anode voltages is chosen. This beam, if **undeflected,** moves down the axis of the cathode-ray tube.

In an electrostatic cathode-ray tube, the anodes perform focusing and acceleration of the cathode beam as previously described. In an electro-magnetic tube, the anodes do not



DEFLECTING PLATES FOR ELECTROSTATIC CATHODE RAY TUBE

A



DIVERGENT DEFLECTING PLATES

B

Figure 3-108

perform focusing. Instead, this is done by a focusing coil wound around the neck of the tube and carrying a current. This current creates a magnetic field inside the tube along the path of the beam and serves to focus it by varying its motion at right angles to its path and to the magnetic field.

5.8.3 Deflecting Elements

Deflection of the focused cathode beam is performed by vertical and horizontal plates in an electrostatic cathode-ray tube and by deflection coils in an electromagnetic cathode-ray tube.

5.8.3.1 Electrostatic Deflection

Two sets of deflecting plates, one for vertical deflection and one for horizontal deflection, are used in an electrostatic tube. Figure 3-108(A), **shows such plates.**

The sharpened beam of electrons comes from d at the left and passes between the two horizontal plates to which voltages can be applied through A and B. If A has a positive voltage and B a negative one, for example, the negatively charged cathode beam will be deflected vertically upward as it passes between these plates. The degree of deflection depends on the strength of the elctrostatic field and on its direction. As the beam then passes between the vertical plates to which are applied voltages through C and D, the beam can be deflected horizontally to the right or left.

S represents the screen in **Figure 3-108(A). The lines od on** S represent horizontal deflection possibilities by the plates CD. The lines ab represent vertical deflection possibilities by the plates AB.

To increase the deflection sensitivity, the plates can be made longer in order to influence the cathode beam for a longer time as it passes between the plates. In such cases, the plates are made divergent at the ends (See Figure 3-108 (B)) to avoid having the beam strike the plates as it is deflected.

5.8.3.2 Electromagnetic Deflection

A typical device used to control the deflection of an electron beam electromagnetically is to use two sets of coils, wound around the neck of the cathode-ray tube and carrying current which produce magnetic fields along the path of the electron beam. One set of coils can produce a vertical field and the other set a horizontal field. Electrons moving in the vertical field will be deflected at right angles to it and to its axial path, that is in a horizontal plane. Electrons moving in a horizontal field directed at right angles to the path of the beam will be deflected in a vertical plane.

5.8.4 Screen

The visual indication of the effects of the cathode-ray beam appears on a target or screen which is placed in a plane substantially at right angles to the axis of the tube. This glass screen is coated on the inside with a phosphor such as calcium tungstate, or zinc silicate. This phosphor becomes luminescent when struck by the beam in any particular spot. The beam can be moved quickly from one spot to an adjacent one, lighting up these successive spots for some small interval. Because of persistence of vision, that is, the ability of the eye senses to retain an image for about a sixteenth of a second after the object is gone, complete screen patterns on the cathode-ray tube can be seen. As long as the electron beam strikes in a given place at least 16 times a second, the spot will appear to the human eye as a source of continuous light with very little flicker.