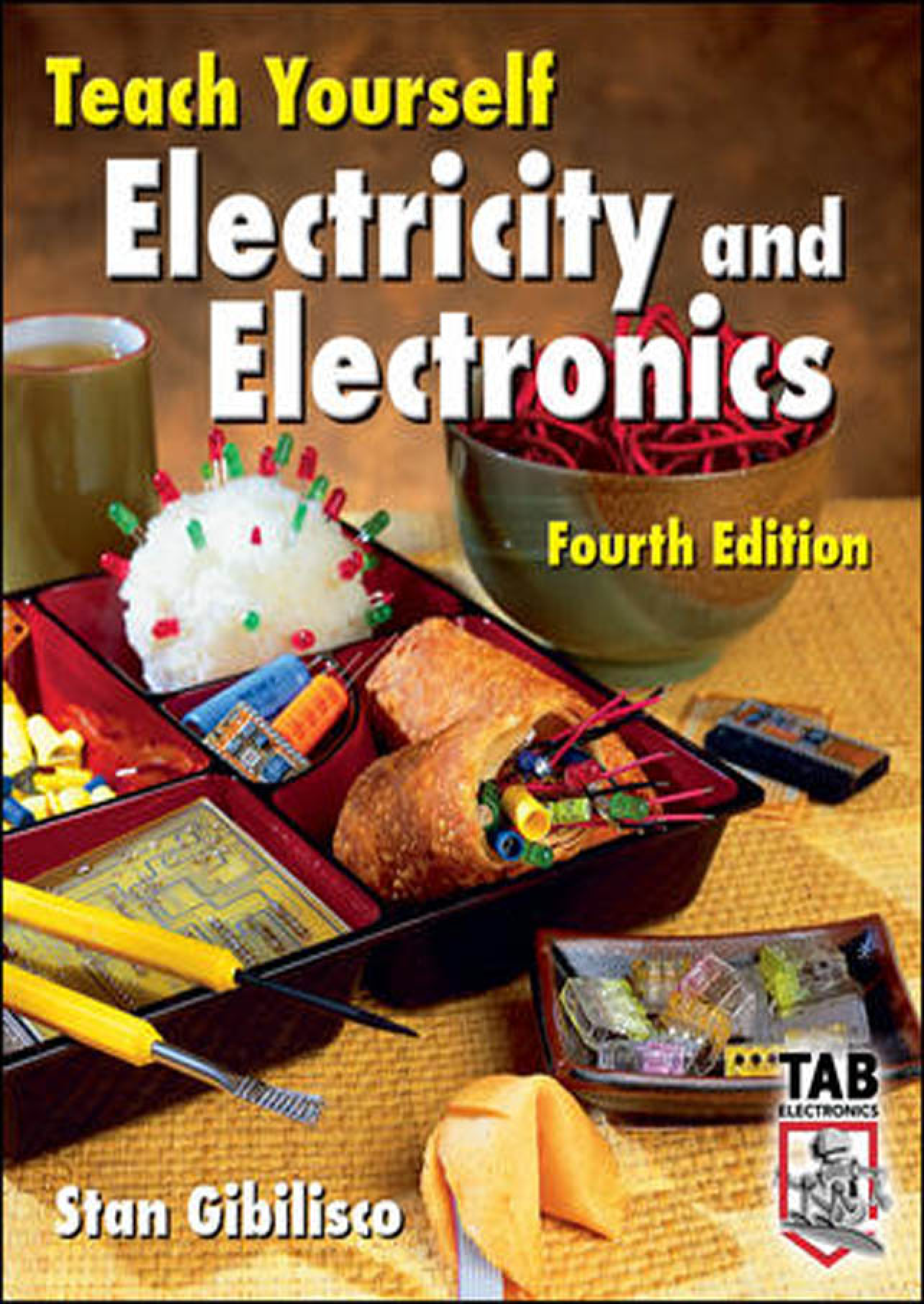
Teach Yourself Electricity and Electronics

Fourth Edition



Stan Gibilisco

Teach Yourself Electricity and Electronics

This page intentionally left blank

Teach Yourself Electricity and Electronics

Fourth Edition

Stan Gibilisco

McGraw-Hill

New York Chicago San Francisco Lisbon London Madrid
Mexico City Milan New Delhi San Juan Seoul
Singapore Sydney Toronto

0-07-148709-3

The material in this eBook also appears in the print version of this title: 0-07-145933-2.

DOI: 10.1036/0071459332

To Tony, Samuel, Tim, Roland, Jack, and Sherri

This page intentionally left blank

Contents

Preface

This book is for people who want to learn the fundamentals of electricity, electronics, and related fields without taking a formal course. The book can also serve as a classroom text. This edition contains new material on transducers, sensors, antennas, monitoring, security, and navigation. Material from previous editions has been updated where appropriate.

As you take this course, you'll encounter hundreds of quiz, test, and exam questions that can help you measure your progress. They are written like the questions found in standardized tests used by educational institutions.

There is a short multiple-choice quiz at the end of every chapter. The quizzes are "open-book." You may refer to the chapter texts when taking them. When you have finished a chapter, take the quiz, write down your answers, and then give your list of answers to a friend. Have the friend tell you your score, but not which questions you got wrong. Because you're allowed to look at the text when taking the quizzes, some of the questions are rather difficult.

At the end of each section, there is a multiple-choice test. These tests are easier than chapter-ending quizzes. Don't look back at the text when taking the tests. A satisfactory score is at least three-quarters of the answers correct.

You will find a final exam at the end of this course. As with the section-ending tests, the questions are not as difficult as those in the chapter-ending quizzes. Don't refer back to the text while taking the final exam. A satisfactory score is at least three-quarters of the answers correct.

The answers to all of the multiple-choice quiz, test, and exam questions are listed in an appendix at the back of this book.

You don't need a mathematical or scientific background for this course. Middle-school algebra, geometry, and physics will suffice. There's no calculus here! I recommend that you complete one chapter a week. That way, in a few months, you'll finish the course. You can then use this book, with its comprehensive index, as a permanent reference.

Suggestions for future editions are welcome.

Stan Gibilisco

This page intentionally left blank

1
PART
Direct Current

This page intentionally left blank

1
CHAPTER
Basic Physical Concepts

IT IS IMPORTANT TO UNDERSTAND SOME SIMPLE, GENERAL PHYSICS PRINCIPLES IN ORDER TO HAVE A full grasp of electricity and electronics. It is not necessary to know high-level mathematics. In science, you can talk about *qualitative* things or *quantitative* things, the "what" versus the "how much." For now, we are concerned only about the "what." The "how much" will come later.

Atoms

All matter is made up of countless tiny particles whizzing around. These particles are extremely dense; matter is mostly empty space. Matter seems continuous because the particles are so small, and they move incredibly fast.

Each chemical *element* has its own unique type of particle, known as its *atom.* Atoms of different elements are always different. The slightest change in an atom can make a tremendous difference in its behavior. You can live by breathing pure oxygen, but you can't live off of pure nitrogen. Oxygen will cause metal to corrode, but nitrogen will not. Wood will burn furiously in an atmosphere of pure oxygen, but will not even ignite in pure nitrogen. Yet both are gases at room temperature and pressure; both are colorless, both are odorless, and both are just about of equal weight. These substances are so different because oxygen has eight *protons*, while nitrogen has only seven. There are many other examples in nature where a tiny change in atomic structure makes a major difference in the way a substance behaves.

Protons, Neutrons, and Atomic Numbers

The part of an atom that gives an element its identity is the *nucleus.* It is made up of two kinds of particles, the *proton* and the *neutron.* These are extremely dense. A teaspoonful of either of these particles, packed tightly together, would weigh tons. Protons and neutrons have just about the same mass, but the proton has an electric charge while the neutron does not.

The simplest element, hydrogen, has a nucleus made up of only one proton; there are usually no neutrons. This is the most common element in the universe. Sometimes a nucleus of hydrogen has a neutron or two along with the proton, but this does not occur very often. These “mutant” forms of hydrogen do, nonetheless, play significant roles in atomic physics.

The second most abundant element is helium. Usually, this atom has a nucleus with two protons and two neutrons. Hydrogen is changed into helium inside the sun, and in the process, energy is given off. This makes the sun shine. The process, called *fusion*, is also responsible for the terrific explosive force of a hydrogen bomb.

Every proton in the universe is just like every other. Neutrons are all alike, too. The number of protons in an element’s nucleus, the *atomic number*, gives that element its identity. The element with three protons is lithium, a light metal that reacts easily with gases such as oxygen or chlorine. The element with four protons is beryllium, also a metal. In general, as the number of protons in an element’s nucleus increases, the number of neutrons also increases. Elements with high atomic numbers, like lead, are therefore much denser than elements with low atomic numbers, like carbon. Perhaps you’ve compared a lead sinker with a piece of coal of similar size, and noticed this difference.

Isotopes and Atomic Weights

For a given element, such as oxygen, the number of neutrons can vary. But no matter what the number of neutrons, the element keeps its identity, based on the atomic number. Differing numbers of neutrons result in various *isotopes* for a given element.

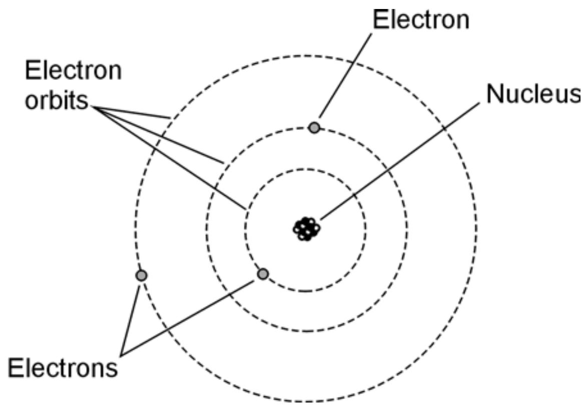
Each element has one particular isotope that is most often found in nature. But all elements have numerous isotopes. Changing the number of neutrons in an element’s nucleus results in a difference in the weight, and also a difference in the density, of the element. Thus, hydrogen containing a neutron or two in the nucleus, along with the proton, is called *heavy hydrogen.*

The *atomic weight* of an element is approximately equal to the sum of the number of protons and the number of neutrons in the nucleus. Common carbon has an atomic weight of about 12, and is called carbon 12 or C12. But sometimes it has an atomic weight of about 14, and is known as carbon 14 or C14.

Electrons

Surrounding the nucleus of an atom are particles having opposite electric charge from the protons. These are the *electrons.* Physicists arbitrarily call the electrons’ charge *negative*, and the protons’ charge *positive.* An electron has exactly the same charge quantity as a proton, but with opposite polarity. The charge on a single electron or proton is the smallest possible electric charge. All charges, no matter how great, are multiples of this unit charge.

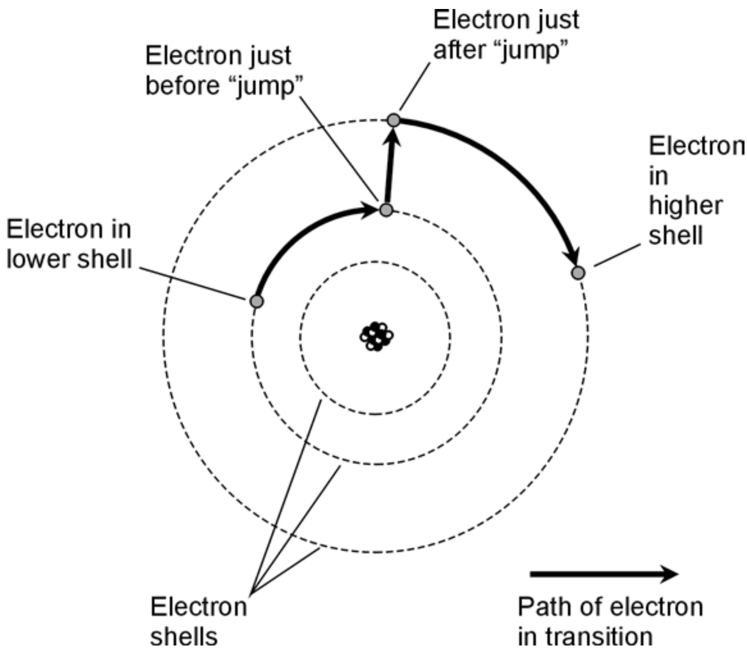
One of the earliest ideas about the atom pictured the electrons embedded in the nucleus, like raisins in a cake. Later, the electrons were seen as orbiting the nucleus, making the atom like a miniature solar system with the electrons as the planets (Fig. 1-1). Still later, this view was modified further. Today, the electrons are seen as so fast-moving, with patterns so complex, that it is not even possible to pinpoint them at any given instant of time. All that can be done is to say that an electron will just as likely be inside a certain sphere as outside. These spheres are known as electron *shells.* Their centers correspond to the position of the atomic nucleus. The farther away from the nucleus the *shell*, the more energy the electron has (Fig. 1-2).



1-1 An early model of the atom, developed around the year 1900, resembled a miniature solar system. The electrons were held in their orbits around the nucleus by electrostatic attraction.

Electrons can move rather easily from one atom to another in some materials. In other substances, it is difficult to get electrons to move. But in any case, it is far easier to move electrons than it is to move protons. Electricity almost always results, in some way, from the motion of electrons in a material. Electrons are much lighter than protons or neutrons. In fact, compared to the nucleus of an atom, the electrons weigh practically nothing.

Generally, the number of electrons in an atom is the same as the number of protons. The negative charges therefore exactly cancel out the positive ones, and the atom is electrically neutral. But



1-2 Electrons move around the nucleus of an atom at defined levels, called *shells*, which correspond to discrete energy states. This is a simplified illustration of an electron gaining energy within an atom.

under some conditions, there can be an excess or shortage of electrons. High levels of radiant energy, extreme heat, or the presence of an electric field (discussed later) can "knock" or "throw" electrons loose from atoms, upsetting the balance.

Ions

If an atom has more or less electrons than protons, that atom acquires an electrical charge. A shortage of electrons results in positive charge; an excess of electrons gives a negative charge. The element's identity remains the same, no matter how great the excess or shortage of electrons. In the extreme case, all the electrons might be removed from an atom, leaving only the nucleus. However, it would still represent the same element as it would if it had all its electrons. A charged atom is called an *ion.* When a substance contains many ions, the material is said to be *ionized.*

A good example of an ionized substance is the atmosphere of the earth at high altitudes. The ultraviolet radiation from the sun, as well as high-speed subatomic particles from space, result in the gases' atoms being stripped of electrons. The ionized gases tend to be found in layers at certain altitudes. These layers are responsible for long-distance radio communications at some frequencies.

Ionized materials generally conduct electricity well, even if the substance is normally not a good conductor. Ionized air makes it possible for a lightning stroke to take place, for example. The ionization, caused by a powerful electric field, occurs along a jagged, narrow channel. After the lightning flash, the nuclei of the atoms quickly attract stray electrons back, and the air becomes electrically neutral again.

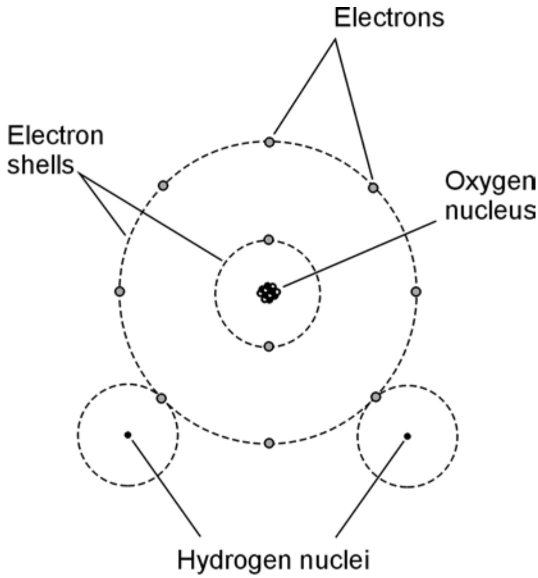
An element might be both an ion and an isotope different from the usual isotope. For example, an atom of carbon might have eight neutrons rather than the usual six, thus being the isotope C14, and it might have been stripped of an electron, giving it a positive unit electric charge and making it an ion.

Compounds

Different elements can join together to share electrons. When this happens, the result is a chemical *compound.* One of the most common compounds is water, the result of two hydrogen atoms joining with an atom of oxygen. There are literally thousands of different chemical compounds that occur in nature.

A compound is different than a simple mixture of elements. If hydrogen and oxygen are mixed, the result is a colorless, odorless gas, just like either element is a gas separately. A spark, however, will cause the molecules to join together; this will liberate energy in the form of light and heat. Under the right conditions, there will be a violent explosion, because the two elements join eagerly. Water is chemically illustrated in Fig. 1-3.

Compounds often, but not always, appear greatly different from any of the elements that make them up. At room temperature and pressure, both hydrogen and oxygen are gases. But water under the same conditions is a liquid. If it gets a few tens of degrees colder, water turns solid at standard pressure. If it gets hot enough, water becomes a gas, odorless and colorless, just like hydrogen or oxygen.



1-3 A simplified diagram of a water molecule. Note the shared electrons.

Another common example of a compound is rust. This forms when iron joins with oxygen. While iron is a dull gray solid and oxygen is a gas, rust is a maroon-red or brownish powder, completely unlike either of the elements from which it is formed.

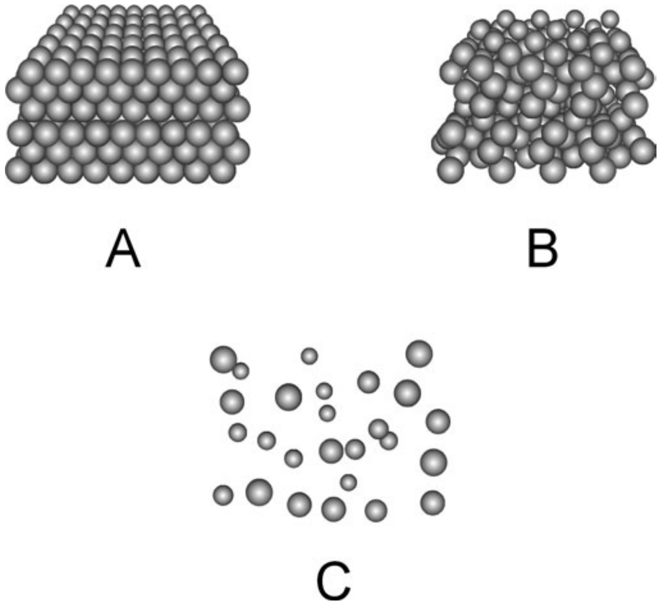
Molecules

When atoms of elements join together to form a compound, the resulting particles are *molecules.* Figure 1-3 is an example of a molecule of water, consisting of three atoms put together.

The natural form of an element is also known as its molecule. Oxygen tends to occur in pairs most of the time in the earth's atmosphere. Thus, an oxygen molecule is sometimes denoted by the symbol O_2. The "O" represents oxygen, and the subscript 2 indicates that there are two atoms per molecule. The water molecule is symbolized H_2O, because there are two atoms of hydrogen and one atom of oxygen in each molecule.

Sometimes oxygen atoms exist all by themselves; then we denote the molecule simply as O. Sometimes there are three atoms of oxygen grouped together. This is the gas called *ozone*, which has received much attention lately in environmental news. It is written O_3.

All matter, whether solid, liquid, or gas, is made of molecules. These particles are always moving. The speed with which they move depends on the temperature. The hotter the temperature, the more rapidly the molecules move around. In a solid, the molecules are interlocked in a sort of rigid pattern, although they vibrate continuously (Fig. 1-4A). In a liquid, they slither and slide around (Fig. 1-4B). In a gas, they rush all over the place, bumping into each other and into solids and liquids adjacent to the gas (Fig. 1-4C).



1-4 Simplified renditions of molecular arrangements in a solid (A), a liquid (B), and a gas (C).

Conductors

In some materials, electrons move easily from atom to atom. In others, the electrons move with difficulty. And in some materials, it is almost impossible to get them to move. An electrical *conductor* is a substance in which the electrons are mobile.

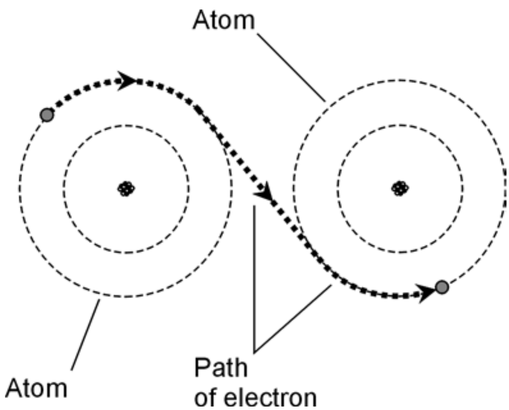
The best conductor at room temperature is pure elemental silver. Copper and aluminum are also excellent electrical conductors. Iron, steel, and various other metals are fair to good conductors of electricity. In most electrical circuits and systems, copper or aluminum wire is used. (Silver is impractical because of its high cost.)

Some liquids are good electrical conductors. Mercury is one example. Salt water is a fair conductor. Gases or mixtures of gases, such as air, are generally poor conductors of electricity. This is because the atoms or molecules are usually too far apart to allow a free exchange of electrons. But if a gas becomes ionized, it can be a fair conductor of electricity.

Electrons in a conductor do not move in a steady stream, like molecules of water through a garden hose. Instead, they are passed from one atom to another right next to it (Fig. 1-5). This happens to countless atoms all the time. As a result, literally trillions of electrons pass a given point each second in a typical electrical circuit.

Insulators

An *insulator* prevents electrical currents from flowing, except occasionally in tiny amounts. Most gases are good electrical insulators. Glass, dry wood, paper, and plastics are other examples. Pure water is a



1-5 In an electrical conductor, certain electrons can pass easily from atom to atom.

good electrical insulator, although it conducts some current with even the slightest impurity. Metal oxides can be good insulators, even though the metal in pure form is a good conductor.

Electrical insulators can be forced to carry current. Ionization can take place; when electrons are stripped away from their atoms, they move more or less freely. Sometimes an insulating material gets charred, or melts down, or gets perforated by a spark. Then its insulating properties are lost, and some electrons flow. An insulating material is sometimes called a *dielectric.* This term arises from the fact that it keeps electrical charges apart, preventing the flow of electrons that would equalize a charge difference between two places. Excellent insulating materials can be used to advantage in certain electrical components such as capacitors, where it is important that electrons not flow.

Porcelain or glass can be used in electrical systems to keep short circuits from occurring. These devices, called insulators, come in various shapes and sizes for different applications. You can see them on high-voltage utility poles and towers. They hold the wire up without running the risk of a short circuit with the tower or a slow discharge through a wet wooden pole.

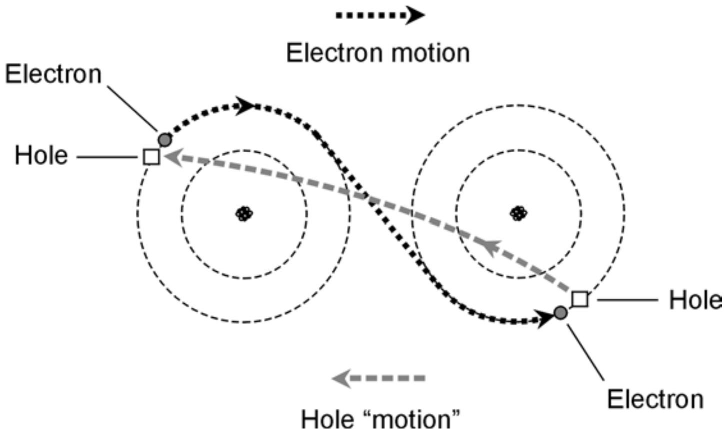
Resistors

Some substances, such as carbon, conduct electricity fairly well but not really well. The conductivity can be changed by adding impurities like clay to a carbon paste, or by winding a thin wire into a coil. Electrical components made in this way are called *resistors.* They are important in electronic circuits because they allow for the control of current flow. The better a resistor conducts, the lower its *resistance;* the worse it conducts, the higher the resistance.

Electrical resistance is measured in units called *ohms.* The higher the value in ohms, the greater the resistance, and the more difficult it becomes for current to flow. For wires, the resistance is sometimes specified in terms of ohms per unit length (foot, meter, kilometer, or mile). In an electrical system, it is usually desirable to have as low a resistance, or ohmic value, as possible. This is because resistance converts electrical energy into heat.

Semiconductors

In a *semiconductor*, electrons flow, but not as well as they do in a conductor. Some semiconductors carry electrons almost as well as good electrical conductors like copper or aluminum; others are almost as bad as insulating materials.



1-6 In a semiconducting material, holes travel in a direction opposite to the direction in which the electrons travel.

Semiconductors are not the same as resistors. In a semiconductor, the material is treated so that it has very special properties.

Semiconductors include certain substances such as silicon, selenium, or gallium, that have been "doped" by the addition of impurities such as indium or antimony. Have you heard of such things as *gallium arsenide*, *metal oxides*, or *silicon rectifiers?* Electrical conduction in these materials is always a result of the motion of electrons. But this can be a quite peculiar movement, and sometimes engineers speak of the movement of *holes* rather than electrons. A hole is a shortage of an electron—you might think of it as a positive ion—and it moves along in a direction opposite to the flow of electrons (Fig. 1-6).

When most of the *charge carriers* are electrons, the semiconductor is called *N-type*, because electrons are negatively charged. When most of the charge carriers are holes, the semiconductor material is known as *P-type* because holes have a positive electric charge. But P-type material does pass some electrons, and N-type material carries some holes. In a semiconductor, the more abundant type of charge carrier is called the *majority carrier.* The less abundant kind is known as the *minority carrier.* Semiconductors are used in *diodes*, *transistors*, and *integrated circuits.* These substances are what make it possible for you to have a computer or a television receiver in a package small enough to hold in your hand.

Current

Whenever there is movement of charge carriers in a substance, there is an electric *current.* Current is measured in terms of the number of electrons or holes passing a single point in 1 second.

A great many charge carriers go past any given point in 1 second, even if the current is small. In a household electric circuit, a 100-watt light bulb draws a current of about *six quintillion* (6 followed by 18 zeros) charge carriers per second. Even the smallest bulb carries *quadrillions* (numbers followed by 15 zeros) of charge carriers every second. It is impractical to speak of a current in terms of charge carriers per second, so it is measured in *coulombs per second* instead. A coulomb is equal to approximately 6,240,000,000,000,000,000 electrons or holes. A current of 1 coulomb per second is called an *ampere*, and this is the standard unit of electric current. A 100-watt bulb in your desk lamp draws about 1 ampere of current.

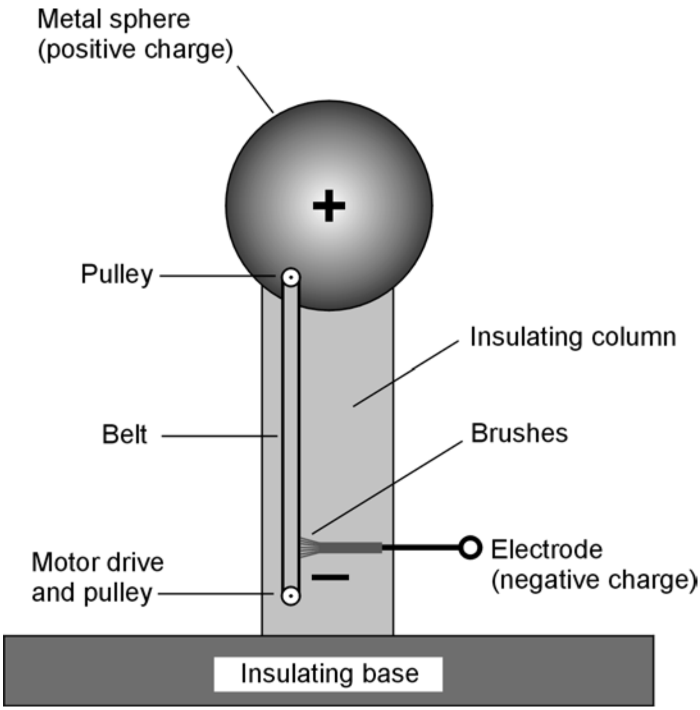
When a current flows through a resistance—and this is always the case because even the best conductors have resistance—heat is generated. Sometimes light and other forms of energy are emitted as well. A light bulb is deliberately designed so that the resistance causes visible light to be generated.

Electric current flows at high speed through any conductor, resistor, or semiconductor. Nevertheless, it is considerably less than the speed of light.

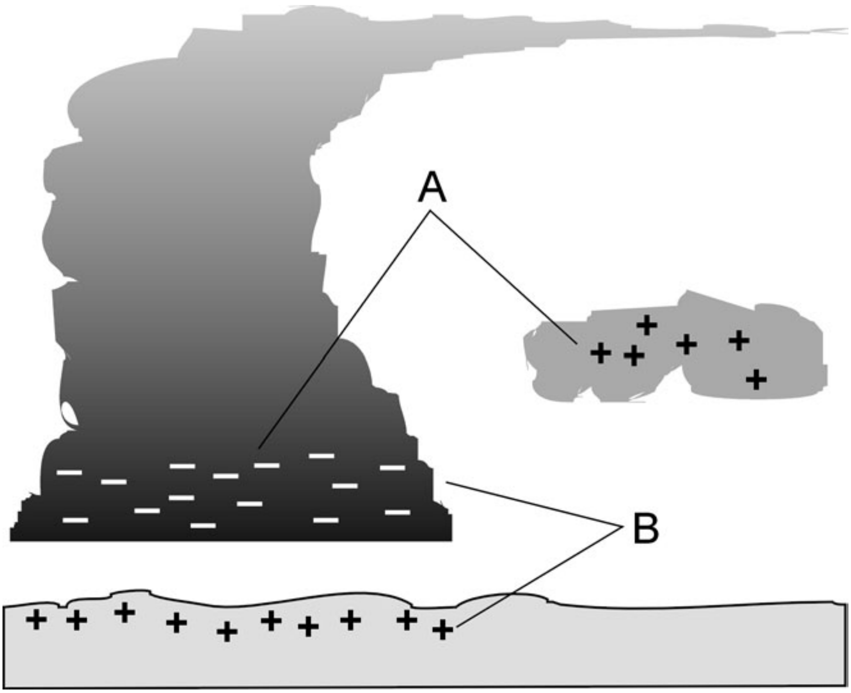
Static Electricity

Charge carriers, particularly electrons, can build up, or become deficient, on things without flowing anywhere. You've experienced this when walking on a carpeted floor during the winter, or in a place where the humidity was low. An excess or shortage of electrons is created on and in your body. You acquire a *charge* of *static electricity.* It's called "static" because it doesn't go anywhere. You don't feel this until you touch some metallic object that is connected to earth ground or to some large fixture; but then there is a *discharge*, accompanied by a spark.

If you were to become much more charged, your hair would stand on end, because every hair would repel every other. Like charges are caused either by an excess or a deficiency of electrons; they repel. The spark might jump an inch, 2 inches, or even 6 inches. Then it would more than startle you; you could get hurt. This doesn't happen with ordinary carpet and shoes, fortunately. But a device called a *Van de Graaff generator*, found in physics labs, can cause a spark this large (Fig. 1-7). Be careful when using this device for physics experiments!



1-7 Simplified illustration of a Van de Graaff generator. This machine can create a charge buildup large enough to produce a spark several centimeters long.



1-8 Electrostatic charges can build up between clouds in a thunderstorm (A), or between a cloud and the surface of the earth (B).

In the extreme, lightning occurs between clouds, and between clouds and ground in the earth's atmosphere. This spark, called a *stroke*, is a magnified version of the spark you get after shuffling around on a carpet. Until the stroke occurs, there is a static charge in the clouds, between different clouds or parts of a cloud, and the ground. In Fig. 1-8, cloud-to-cloud (A) and cloud-to-ground (B) static buildups are shown. In the case at B, the positive charge in the earth follows along beneath the storm cloud. The current in a lightning stroke is usually several tens of thousands, or hundreds of thousands, of amperes. But it takes place only for a fraction of a second. Still, many coulombs of charge are displaced in a single bolt of lightning.

Electromotive Force

Current can only flow if it gets a "push." This can be caused by a buildup of static electric charges, as in the case of a lightning stroke. When the charge builds up, with positive polarity (shortage of electrons) in one place and negative polarity (excess of electrons) in another place, a powerful *electromotive force* (EMF) exists. This force is measured in units called *volts.*

Ordinary household electricity has an effective voltage of between 110 and 130; usually it is about 117. A car battery has an EMF of 12 to 14 volts. The static charge that you acquire when walking on a carpet with hard-soled shoes is often several thousand volts. Before a discharge of lightning, millions of volts exist. An EMF of 1 volt, across a resistance of 1 ohm, will cause a current of 1 ampere to flow. This is a classic relationship in electricity, and is stated generally as *Ohm's Law.* If the EMF is doubled, the current is doubled. If the resistance is doubled, the current is cut in half. This important law of electrical circuit behavior is covered in detail later in this book.

It is possible to have an EMF without having any current. This is the case just before a lightning stroke occurs, and before you touch a metal object after walking on a carpet. It is also true between the two wires of an electric lamp when the switch is turned off. It is true of a dry cell when there is nothing connected to it. There is no current, but a current is possible given a conductive path between the two points. Voltage, or EMF, is sometimes called *potential* or *potential difference* for this reason.

Even a huge EMF does not necessarily drive much current through a conductor or resistance. A good example is your body after walking around on the carpet. Although the voltage seems deadly in terms of numbers (thousands), there are not many coulombs of static-electric charge that can accumulate on an object the size of your body. Therefore, in relative terms, not that many electrons flow through your finger when you touch a radiator. This is why you don't get a severe shock.

If there are plenty of coulombs available, a small voltage, such as 117 volts (or even less) can cause a lethal current. This is why it is dangerous to repair an electrical device with the power on. The power plant will pump an unlimited number of coulombs of charge through your body if you are not careful.

Nonelectrical Energy

In electricity and electronics, there are phenomena that involve other forms of energy besides electrical energy. Visible light is an example. A light bulb converts electricity into radiant energy that you can see. This was one of the major motivations for people like Thomas Edison to work with electricity. Visible light can also be converted into electric current or voltage. A *photovoltaic cell* does this.

Light bulbs always give off some heat, as well as visible light. Incandescent lamps actually give off more energy as heat than as light. You are certainly acquainted with electric heaters, designed for the purpose of changing electricity into heat energy. This heat is a form of radiant energy called *infrared* (IR). It is similar to visible light, except that the waves are longer and you can't see them.

Electricity can be converted into other radiant-energy forms, such as *radio waves*, *ultraviolet* (UV), and *X rays.* This is done by specialized devices such as radio transmitters, sunlamps, and electron tubes. Fast-moving protons, neutrons, electrons, and atomic nuclei are an important form of energy. The energy from these particles is sometimes sufficient to split atoms apart. This effect makes it possible to build an atomic reactor whose energy can be used to generate electricity.

When a conductor moves in a magnetic field, electric current flows in that conductor. In this way, mechanical energy is converted into electricity. This is how an *electric generator* works. Generators can also work backward. Then you have a *motor* that changes electricity into useful mechanical energy.

A magnetic field contains energy of a unique kind. The science of *magnetism* is closely related to electricity. Magnetic phenomena are of great significance in electronics. The oldest and most universal source of magnetism is the *geomagnetic field* surrounding the earth, caused by alignment of iron atoms in the core of the planet.

A changing magnetic field creates a fluctuating electric field, and a fluctuating electric field produces a changing magnetic field. This phenomenon, called *electromagnetism*, makes it possible to send wireless signals over long distances. The electric and magnetic fields keep producing one another over and over again through space.

Chemical energy is converted into electricity in *dry cells*, *wet cells*, and *batteries*. Your car battery is an excellent example. The acid reacts with the metal electrodes to generate an electromotive force. When the two poles of the batteries are connected, current results. The chemical reaction continues, keeping the current going for a while. But the battery can only store a certain amount of chemical energy. Then it "runs out of juice," and the supply of chemical energy must be restored by *charging*. Some cells and batteries, such as lead-acid car batteries, can be recharged by driving current through them, and others, such as most flashlight and transistor-radio batteries, cannot.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct answers out of these 20 questions. The answers are listed in the back of this book.

1. The atomic number of an element is determined by
   (a) the number of neutrons.
   (b) the number of protons.
   (c) the number of neutrons plus the number of protons.
   (d) the number of electrons.

2. The atomic weight of an element is approximately determined by
   (a) the number of neutrons.
   (b) the number of protons.
   (c) the number of neutrons plus the number of protons.
   (d) the number of electrons.

3. Suppose there is an atom of oxygen, containing eight protons and eight neutrons in the nucleus, and two neutrons are added to the nucleus. What is the resulting atomic weight?
   (a) 8
   (b) 10
   (c) 16
   (d) 18

4. An ion
   (a) is electrically neutral.
   (b) has positive electric charge.
   (c) has negative electric charge.
   (d) can have either a positive or negative charge.

5. An isotope
   (a) is electrically neutral.
   (b) has positive electric charge.
   (c) has negative electric charge.
   (d) can have either a positive or negative charge.

6. A molecule
   (a) can consist of a single atom of an element.
   (b) always contains two or more elements.
   (c) always has two or more atoms.
   (d) is always electrically charged.
7. In a compound,
   (a) there can be a single atom of an element.
   (b) there must always be two or more elements.
   (c) the atoms are mixed in with each other but not joined.
   (d) there is always a shortage of electrons.
8. An electrical insulator can be made a conductor
   (a) by heating it.
   (b) by cooling it.
   (c) by ionizing it.
   (d) by oxidizing it.
9. Of the following substances, the worst conductor is
   (a) air.
   (b) copper.
   (c) iron.
   (d) salt water.
10. Of the following substances, the best conductor is
    (a) air.
    (b) copper.
    (c) iron.
    (d) salt water.
11. Movement of holes in a semiconductor
    (a) is like a flow of electrons in the same direction.
    (b) is possible only if the current is high enough.
    (c) results in a certain amount of electric current.
    (d) causes the material to stop conducting.
12. If a material has low resistance, then
    (a) it is a good conductor.
    (b) it is a poor conductor.
    (c) the current flows mainly in the form of holes.
    (d) current can flow only in one direction.
13. A coulomb
    (a) represents a current of 1 ampere.
    (b) flows through a 100-watt light bulb.
    (c) is equivalent to 1 ampere per second.
    (d) is an extremely large number of charge carriers.

14. A stroke of lightning
    (a) is caused by a movement of holes in an insulator.
    (b) has a very low current.
    (c) is a discharge of static electricity.
    (d) builds up between clouds.
15. The volt is the standard unit of
    (a) current.
    (b) charge.
    (c) electromotive force.
    (d) resistance.
16. If an EMF of 1 volt is placed across a resistance of 2 ohms, then the current is
    (a) half an ampere.
    (b) 1 ampere.
    (c) 2 amperes.
    (d) impossible to determine.
17. A backward-working electric motor, in which mechanical rotation is converted to electricity, is best described as
    (a) an inefficient, energy-wasting device.
    (b) a motor with the voltage connected the wrong way.
    (c) an electric generator.
    (d) a magnetic field.
18. In a battery, chemical energy can sometimes be replenished by
    (a) connecting it to a light bulb.
    (b) charging it.
    (c) discharging it.
    (d) no means known; when a battery is dead, you must throw it away.
19. A fluctuating magnetic field
    (a) produces an electric current in an insulator.
    (b) magnetizes the earth.
    (c) produces a fluctuating electric field.
    (d) results from a steady electric current.
20. Visible light is converted into electricity
    (a) in a dry cell.
    (b) in a wet cell.
    (c) in an incandescent bulb.
    (d) in a photovoltaic cell.

2
CHAPTER
Electrical Units

THIS CHAPTER EXPLAINS, IN MORE DETAIL, STANDARD UNITS THAT DEFINE THE BEHAVIOR OF DIRECT-current (dc) circuits. Many of these rules also apply to utility alternating-current (ac) circuits.

The Volt

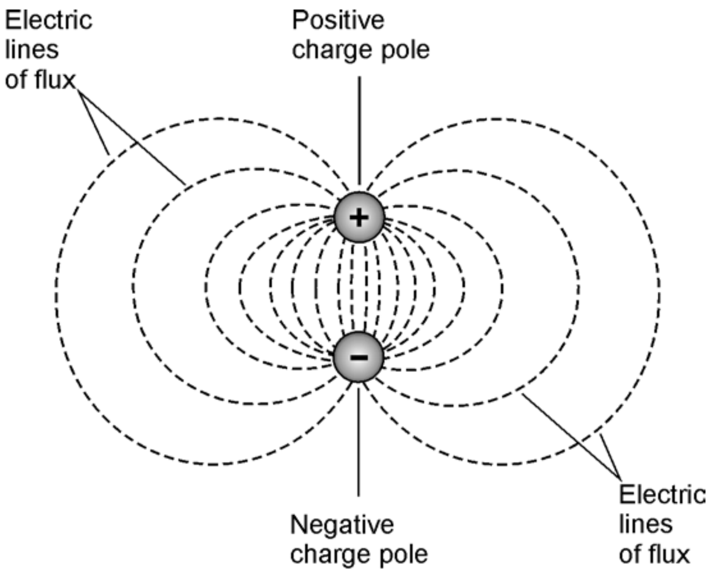
In Chap. 1, you learned a little about the volt, the standard unit of electromotive force (EMF) or *potential difference.*

An accumulation of electrostatic charge, such as an excess or shortage of electrons, is always associated with a voltage. There are other situations in which voltages exist. Voltage can be generated at a power plant, produced in an electrochemical reaction, or caused by light rays striking a semiconductor chip. It can be produced when an object is moved in a magnetic field, or is placed in a fluctuating magnetic field.

A potential difference between two points produces an *electric field*, represented by *electric lines of flux* (Fig. 2-1). There is a pole that is relatively positive, with fewer electrons, and one that is relatively negative, with more electrons. The positive pole does not necessarily have a deficiency of electrons compared with neutral objects, and the negative pole does not always have a surplus of electrons relative to neutral objects. But the negative pole always has more electrons than the positive pole.

The abbreviation for volt (or volts) is V. Sometimes, smaller units are used. The *millivolt* (mV) is equal to a thousandth (0.001) of a volt. The *microvolt* (μV) is equal to a millionth (0.000001) of a volt. It is sometimes necessary to use units larger than the volt. One *kilovolt* (kV) is one thousand volts (1000 V). One *megavolt* (MV) is 1 million volts (1,000,000 V) or one thousand kilovolts (1000 kV).

In a dry cell, the voltage is usually between 1.2 and 1.7 V; in a car battery, it is 12 to 14 V. In household utility wiring, it is a low-frequency alternating current of about 117 V for electric lights and most appliances, and 234 V for a washing machine, dryer, oven, or stove. In television sets, transformers convert 117 V to around 450 V for the operation of the picture tube. In some broadcast transmitters, the voltage can be several kilovolts.



2-1 Electric lines of flux always exist near poles of electric charge.

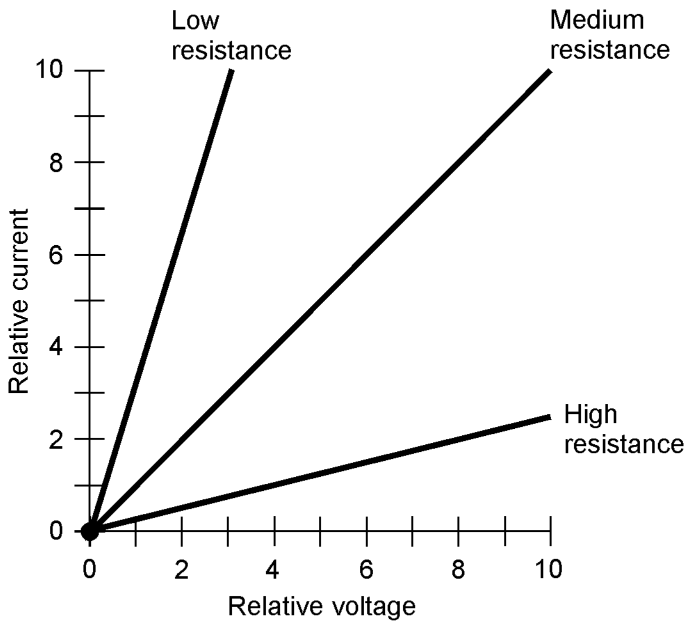
The largest voltages on our planet occur between clouds, or between clouds and the ground, in thundershowers. This potential difference can build up to several tens of megavolts. The existence of a voltage always means that *charge carriers*, which are electrons in a conventional circuit, flow between two points if a conductive path is provided. Voltage represents the driving force that impels charge carriers to move. If all other factors are held constant, high voltages produce a faster flow of charge carriers, and therefore larger currents, than low voltages. But that's an oversimplification in most real-life scenarios, where other factors are hardly ever constant!

Current Flow

If a conducting or semiconducting path is provided between two poles having a potential difference, charge carriers flow in an attempt to equalize the charge between the poles. This flow of *current* continues as long as the path is provided, and as long as there is a charge difference between the poles.

Sometimes the charge difference is equalized after a short while. This is the case, for example, when you touch a radiator after shuffling around on the carpet while wearing hard-soled shoes. It is also true in a lightning stroke. In these instances, the charge is equalized in a fraction of a second. In other cases, the charge takes longer to be used up. This happens if you short-circuit a dry cell. Within a few minutes, the cell "runs out of juice" if you put a wire between the positive and negative terminals. If you put a bulb across the cell, say with a flashlight, it takes an hour or two for the charge difference to drop to zero.

In household electric circuits, the charge difference is never equalized, unless there's a power failure. Of course, if you short-circuit an outlet (don't!), the fuse or breaker will blow or trip, and the charge difference will immediately drop to zero. But if you put a 100-watt bulb at the outlet, the charge difference will be maintained as the current flows. The power plant can keep a potential difference across a lot of light bulbs indefinitely.



2-2 Relative current as a function of relative voltage for low, medium, and high resistances.

Have you heard that it is current, not voltage, that kills? This is a literal truth, but it plays on semantics. It's like saying "It's the heat, not the fire, that burns you." Naturally! But there can only be a deadly current if there is enough voltage to drive it through your body. You don't have to worry when handling flashlight cells, but you'd better be extremely careful around household utility circuits. A voltage of 1.2 to 1.7 V can't normally pump a dangerous current through you, but a voltage of 117 V almost always can.

In an electric circuit that always conducts equally well, the current is directly proportional to the applied voltage. If you double the voltage, you double the current. If the voltage is cut in half, the current is cut in half too. Figure 2-2 shows this relationship as a graph in general terms. It assumes that the power supply can provide the necessary number of charge carriers.

The Ampere

Current is a measure of the rate at which charge carriers flow. The standard unit is the *ampere.* This represents one coulomb (6,240,000,000,000,000,000) of charge carriers flowing every second past a given point.

An ampere is a comparatively large amount of current. The abbreviation is A. Often, current is specified in terms of *milliamperes*, abbreviated mA, where 1 mA = 0.001 A, or a thousandth of an ampere. You will also sometimes hear of *microamperes* (μA), where 1 μA = 0.000001 A or 0.001 mA, which is a millionth of an ampere. It is increasingly common to hear about *nanoamperes* (nA), where 1 nA = 0.001 μA = 0.000000001 A, which is a thousandth of a millionth of an ampere.

A current of a few milliamperes will give you a startling shock. About 50 mA will jolt you severely, and 100 mA can cause death if it flows through your chest cavity. An ordinary 100-watt light bulb draws about 1 A of current in a household utility circuit. An electric iron draws approximately

10 A; an entire household normally uses between 10 and 50 A, depending on the size of the house and the kinds of appliances it has, and also on the time of day, week, or year.

The amount of current that flows in an electrical circuit depends on the voltage, and also on the *resistance.* There are some circuits in which extremely large currents, say 1000 A, can flow. This will happen through a metal bar placed directly at the output of a massive electric generator. The resistance is extremely low in this case, and the generator is capable of driving huge numbers of charge carriers through the bar every second. In some semiconductor electronic devices, such as microcomputers, a few nanoamperes will suffice for many complicated processes. Some electronic clocks draw so little current that their batteries last as long as they would if left on the shelf without being put to any use.

Resistance and the Ohm

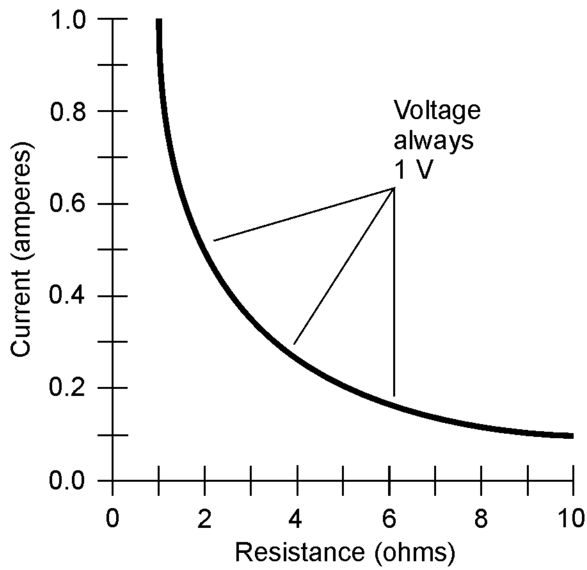
Resistance is a measure of the opposition that a circuit offers to the flow of electric current. You can compare it to the diameter of a hose. In fact, for metal wire, this is an excellent analogy: small-diameter wire has high resistance (a lot of opposition to current), and large-diameter wire has low resistance (not much opposition to current). The type of metal makes a difference too. For example, steel wire has higher resistance for a given diameter than copper wire.

The standard unit of resistance is the *ohm.* This is sometimes symbolized by the uppercase Greek letter omega (Ω). You'll sometimes hear about *kilohms* (symbolized k or kΩ), where 1 kΩ = 1000 Ω, or about *megohms* (symbolized M or MΩ), where 1 MΩ = 1000 kΩ = 1,000,000 Ω.

Electric wire is sometimes rated for *resistivity.* The standard unit for this purpose is the *ohm per foot* (ohm/ft or Ω/ft) or the *ohm per meter* (ohm/m or Ω/m). You might also come across the unit *ohm per kilometer* (ohm/km or Ω/km). Table 2-1 shows the resistivity for various common sizes of solid copper wire at room temperature, as a function of the wire size as defined by a scheme known as the *American Wire Gauge* (AWG).

Table 2-1. Approximate resistivity at room temperature for solid copper wire as a function of the wire size in American Wire Gauge (AWG).

Wire size, AWG #	Resistivity, ohms/km
2	0.52
4	0.83
6	1.3
8	2.7
10	3.3
12	5.3
14	8.4
16	13
18	21
20	34
22	54
24	86
26	140
28	220
30	350

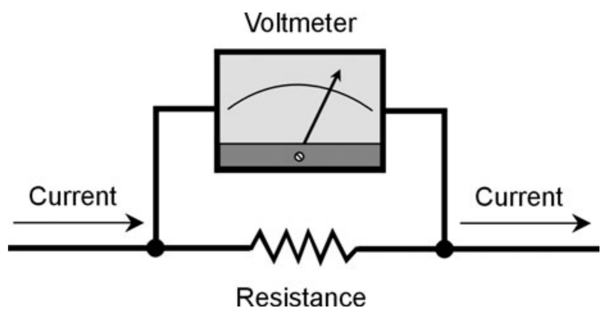


2-3 Current as a function of resistance through an electric device for a constant voltage of 1 V.

When 1 V is placed across 1 Ω of resistance, assuming that the power supply can deliver an unlimited number of charge carriers, there is a current of 1 A. If the resistance is doubled to 2 Ω, the current decreases to 0.5 A. If the resistance is cut by a factor of 5 to 0.2 Ω, the current increases by the same factor, to 5 A. The current flow, for a constant voltage, is said to be *inversely proportional* to the resistance. Figure 2-3 is a graph that shows various currents, through various resistances, given a constant voltage of 1 V across the whole resistance.

Resistance has another property. If there is a current flowing through a resistive material, there is always a potential difference across the resistive component (called a *resistor*). This is shown in Fig. 2-4. In general, this voltage is directly proportional to the current through the resistor. This behavior of resistors is useful in the design of electronic circuits, as you will learn later in this book.

Electrical circuits always have some resistance. There is no such thing as a perfect conductor. When some metals are chilled to temperatures near *absolute zero*, they lose practically all of their resistance, but they never become absolutely perfect, resistance-free conductors. This phenomenon, about which you might have heard, is called *superconductivity.*



2-4 Whenever current passes through a component having resistance, a voltage exists across that component.

Just as there is no such thing as a perfectly resistance-free substance, there isn't a truly infinite resistance, either. Even air conducts to some extent, although the effect is usually so small that it can be ignored. In some electronic applications, materials are selected on the basis of how "nearly infinite" their resistance is.

In electronics, the resistance of a component often varies, depending on the conditions under which it is operated. A transistor, for example, might have high resistance some of the time, and low resistance at other times. High/low resistance variations can be made to take place thousands, millions, or billions of times each second. In this way, oscillators, amplifiers, and digital devices function in radio receivers and transmitters, telephone networks, digital computers, and satellite links (to name just a few applications).

Conductance and the Siemens

Electricians and electrical engineers sometimes talk about the *conductance* of a material, rather than about its resistance. The standard unit of conductance is the *siemens*, abbreviated S. When a component has a conductance of 1 S, its resistance is 1 Ω. If the resistance is doubled, the conductance is cut in half, and vice versa. Therefore, conductance is the reciprocal of resistance.

If you know the resistance of a component or circuit in ohms, you can get the conductance in siemens: divide 1 by the resistance. If you know the conductance in siemens, you can get the resistance: divide 1 by the conductance. Resistance, as a variable quantity, is denoted by an italicized, uppercase letter R. Conductance, as a variable quantity, is denoted as an italicized, uppercase letter G. If we express R in ohms and G in siemens, then the following two equations describe their relationship:

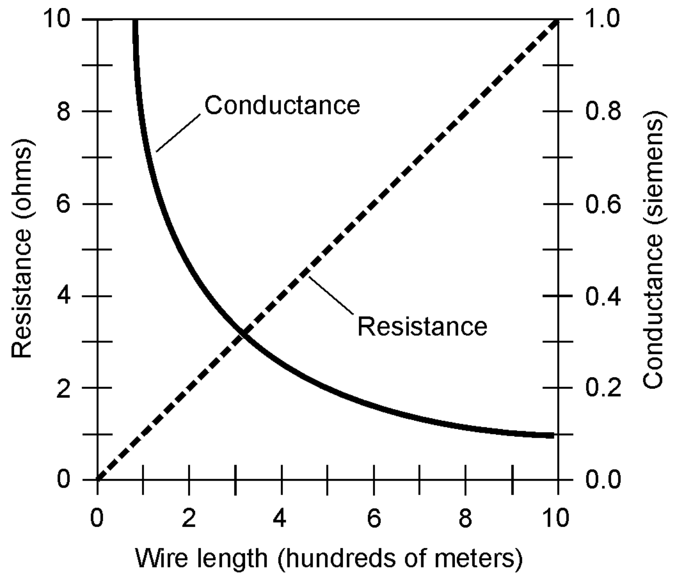
$$G = 1/R$$
$$R = 1/G$$

Units of conductance much smaller than the siemens are often used. A resistance of 1 kΩ is equal to 1 *millisiemens* (1 mS). If the resistance is 1 MΩ, the conductance is one *microsiemens* (1 μS). You'll sometimes hear about *kilosiemens* (kS) or *megasiemens* (MS), representing resistances of 0.001 Ω and 0.000001 Ω (a thousandth of an ohm and a millionth of an ohm, respectively). Short lengths of heavy wire have conductance values in the range of kilosiemens. Heavy metal rods can have conductances in the megasiemens range.

Suppose a component has a resistance of 50 Ω. Then its conductance, in siemens, is 1/50 S, which is equal to 0.02 S. We can call this 20 mS. Or imagine a piece of wire with a conductance of 20 S. Its resistance is 1/20 Ω, which is equal to 0.05 Ω. You will not often hear the term *milliohm.* But you could say that this wire has a resistance of 50 mΩ, and you would be technically right.

Determining *conductivity* is tricky. If wire has a resistivity of 10 Ω/km, you can't say that it has a conductivity of 1/10, or 0.1, S/km. It is true that a kilometer of such wire has a conductance of 0.1 S, but 2 km of the wire has a resistance of 20 Ω (because there is twice as much wire). That is not twice the conductance, but half. If you say that the conductivity of the wire is 0.1 S/km, then you might be tempted to say that 2 km of the wire has 0.2 S of conductance. That would be a mistake! Conductance decreases with increasing wire length.

Figure 2-5 illustrates the resistance and conductance values for various lengths of wire having a resistivity of 10 Ω/km.



2-5 Resistance and conductance for various lengths of wire having a resistivity of 10 ohms per kilometer.

Power and the Watt

Whenever current flows through a resistance, heat results. The heat can be measured in *watts* (symbolized W) and represents electrical *power.* (As a variable quantity in equations, power is denoted by the uppercase italic letter *P*.) Power can be manifested in many forms, such as mechanical motion, radio waves, visible light, or noise. But heat is always present, in addition to any other form of power, in an electrical or electronic device. This is because no equipment is 100 percent efficient. Some power always goes to waste, and this waste is almost all in the form of heat.

Look again at Fig. 2-4. There is a certain voltage across the resistor, not specifically indicated. There's also a current flowing through the resistance, and it is not quantified in the diagram, either. Suppose we call the voltage E and the current I, in volts (V) and amperes (A), respectively. Then the power in watts dissipated by the resistance, call it P, is the product of the voltage in volts and the current in amperes:

$$P = EI$$

If the voltage E across the resistance is caused by two flashlight cells in series, giving 3 V, and if the current I through the resistance (a light bulb, perhaps) is 0.1 A, then $E = 3$ V and $I = 0.1$ A, and we can calculate the power P in watts as follows:

$$P = EI = 3 \times 0.1 = 0.3 \text{ W}$$

Suppose the voltage is 117 V, and the current is 855 mA. To calculate the power, we must convert the current into amperes: 855 mA = 855/1000 A = 0.855 A. Then:

$$P = EI = 117 \times 0.855 = 100 \text{ W}$$

Table 2-2. Prefix multipliers from 0.000000000001 (trillionths, or units of 10^{-12}) to 1,000,000,000,000 (trillions, or units of 10^{12}).

Prefix	Symbol	Multiplier
pico-	p	0.000000000001 (or 10^{-12})
nano-	n	0.000000001 (or 10^{-9})
micro-	μ	0.000001 (or 10^{-6})
milli-	m	0.001 (or 10^{-3})
kilo-	k	1000 (or 10^{3})
mega-	M	1,000,000 (or 10^{6})
giga-	G	1,000,000,000 (or 10^{9})
tera-	T	1,000,000,000,000 (or 10^{12})

You will often hear about *milliwatts* (mW), *microwatts* (μW), *kilowatts* (kW), and *megawatts* (MW). By now, you should be able to tell from the prefixes what these units represent. Otherwise, you can refer to Table 2-2. This table lists the most commonly used *prefix multipliers* in electricity and electronics.

Sometimes you need to use the power equation to find currents or voltages. Then you should use $I = P/E$ to find current, or $E = P/I$ to find voltage. Always remember to convert, if necessary, to the standard units of volts, amperes, and watts before performing the calculations.

A Word about Notation

Have you noticed some strange things about the notation yet? If you're observant, you have! Why, you might ask, are italics sometimes used, and sometimes not used? Something should be said early in this course about notation, because it can get confusing with all the different symbols and abbreviations. Sometimes, symbols and abbreviations appear in italics, and sometimes they do not. You'll see subscripts often, and sometimes even they are italicized! Here are some rules that apply to notation in electricity and electronics:

- Symbols for specific units, such as volts, amperes, and ohms, are not italicized.
- Symbols for objects or components, such as resistors, batteries, and meters, are not italicized.
- Quantifying prefixes, such as "kilo-" or "micro-," are not italicized.
- Labeled points in drawings might or might not be italicized; it doesn't matter as long as a diagram is consistent with itself.
- Symbols for mathematical constants and variables, such as time, are italicized.
- Symbols for electrical quantities, such as voltage, current, resistance, and power, are italicized.
- Symbols and abbreviations for modifiers might or might not be italicized; it doesn't matter as long as a document is consistent with itself.
- Numeric subscripts are not italicized.
- For nonnumeric subscripts, the same rules apply as for general symbols.

Some examples are R (not italicized) for resistor, *R* (italicized) for resistance, *P* (italicized) for power, W (not italicized) for watts, V (not italicized) for volts, *E* or *V* (italicized) for voltage, A (not italicized) for amperes, *I* (italicized) for current, *f* (italicized) for frequency, and *t* (italicized) for time.

Once in a while you will see the same symbol italicized in one place and not in another—in the same circuit diagram or discussion! We might, for example, talk about "resistor number 3" (symbolized R_3), and then later in the same paragraph talk about its value as "resistance number 3" (Symbolized R_3). Still later we might talk about "the *n*th resistor in a series connection" (R_n) and then "the *n*th resistance in a series combination of resistances" (R_n).

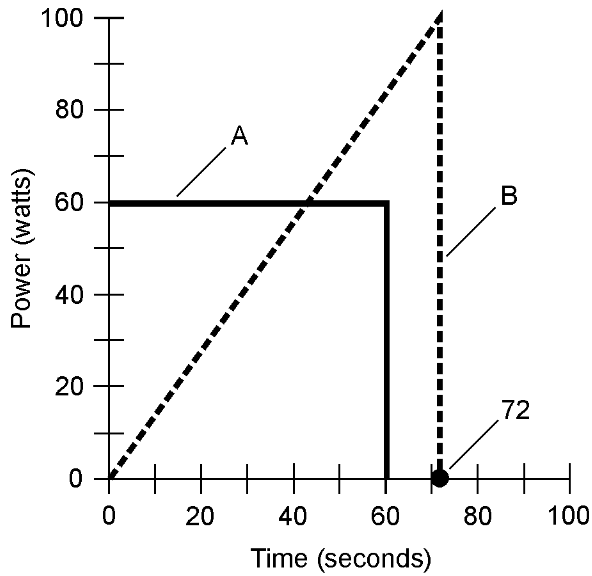
These differences in notation, while subtle (and, some people will say, picayune) are followed in this book, and they are pretty much agreed upon by convention. They are important because they tell the reader exactly what a symbol stands for in a diagram, discussion, or mathematical equation. "Resistor" and "resistance" are vastly different things—as different from each other as a garden hose (the object) and the extent to which it impedes the flow of water (the phenomenon). With this in mind, let us proceed!

Energy and the Watt-Hour

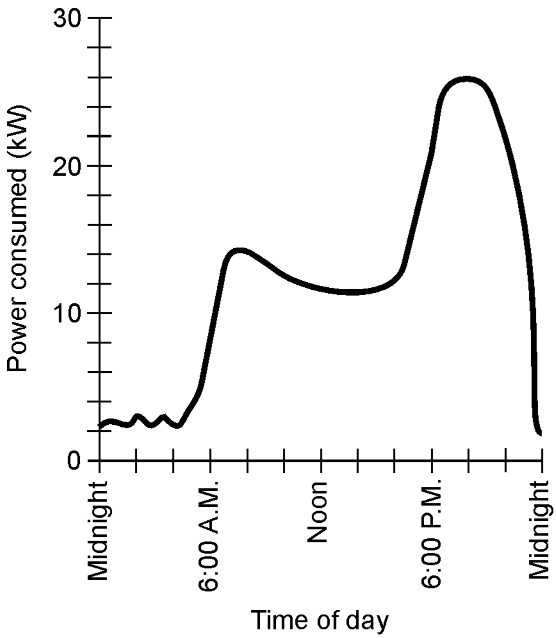
Have you heard the terms "power" and "energy" used interchangeably, as if they mean the same thing? They don't! *Energy* is power dissipated over a length of time. *Power* is the rate at which energy is expended. Physicists measure energy in units called *joules.* One joule (1 J) is the equivalent of a *watt-second*, which is the equivalent of 1 watt of power dissipated for 1 second of time (1 W · s or Ws). In electricity, you'll more often encounter the *watt-hour* (symbolized W · h or Wh) or the *kilowatt-hour* (symbolized kW · h or kWh). As their names imply, a watt-hour is the equivalent of 1 W dissipated for 1 h, and 1 kWh is the equivalent of 1 kW of power dissipated for 1 h.

A watt-hour of energy can be dissipated in an infinite number of different ways. A 60-W bulb consumes 60 Wh in 1 h, the equivalent of a watt-hour per minute (1 Wh/min). A 100-W bulb consumes 1 Wh in 1/100 h, or 36 s. Besides these differences, the rate of power dissipation in real-life circuits often changes with time. This can make the determination of consumed energy complicated, indeed.

Figure 2-6 illustrates two hypothetical devices that consume 1 Wh of energy. Device A uses its power at a constant rate of 60 W, so it consumes 1 Wh in 1 min. The power consumption rate of



2-6 Two devices that burn 1 Wh of energy. Device A dissipates a constant amount of power. Device B dissipates a variable amount of power.



2-7 A graph showing the amount of power consumed by a hypothetical household, as a function of the time of day.

device B varies, starting at zero and ending up at quite a lot more than 60 W. How do you know that this second device really consumes 1 Wh of energy? You must determine the area under the curve in the graph. In this case, figuring out this area is easy, because the enclosed object is a triangle. The area of a triangle is equal to half the product of the base length and the height. Device B is powered up for 72 s, or 1.2 min; this is 1.2/60 = 0.02 h. Then the area under the curve is $1/2 \times 100 \times 0.02 = 1$ Wh.

When calculating energy values, you must always remember the units you're using. In this case the unit is the watt-hour, so you must multiply watts by hours. If you multiply watts by minutes, or watts by seconds, you'll get the wrong kind of units in your answer.

Often, the curves in graphs like these are complicated. Consider the graph of power consumption in your home, versus time, for a day. It might look like the curve in Fig. 2-7. Finding the area under this curve is not easy. But there is another way to determine the total energy burned by your household over a period of time. That is by means of a meter that measures electrical energy in kilowatt-hours. Every month, without fail, the power company sends its representative to read your electric meter. This person takes down the number of kilowatt-hours displayed, subtracts the number from the reading taken the previous month, and a few days later you get a bill. This meter automatically keeps track of total consumed energy, without anybody having to go through high-level mathematical calculations to find the areas under irregular curves such as the graph of Fig. 2-7.

Other Energy Units

The joule, while standard among scientists, is not the only energy unit in existence! Another unit is the *erg*, equivalent to one ten-millionth (0.0000001) of a joule. The erg is used in lab experiments involving small amounts of expended energy.

Table 2-3. Conversion factors between joules and various other energy units.

Unit	To convert energy in this unit to energy in joules, multiply by	To convert energy in joules to energy in this unit, multiply by
British thermal units (Btu)	1055	0.000948
Electron volts (eV)	1.6×10^{-19}	6.2×10^{18}
Ergs	0.0000001 (or 10^{-7})	10,000,000 (or 10^{7})
Foot-pounds (ft-lb)	1.356	0.738
Watt-hours (Wh)	3600	0.000278
Kilowatt-hours (kWh)	3,600,000 (or 3.6×10^{6})	0.000000278 (or 2.78×10^{-7})

Most folks have heard of the *British thermal unit* (Btu), equivalent to 1055 joules. This is the energy unit commonly used to define the cooling or heating capacity of air-conditioning equipment. To cool your room from 85 to 78°F needs a certain amount of energy, perhaps best specified in Btu. If you are getting an air conditioner or furnace installed in your home, an expert will come look at your situation, and determine the size of air-conditioning/heating unit that best suits your needs. That person will likely tell you how powerful the unit should be in terms of its ability to heat or cool in Btu per hour (Btu/h).

Physicists also use, in addition to the joule, a unit of energy called the *electron volt* (eV). This is a tiny unit of energy, equal to just 0.00000000000000000016 joule (there are 18 zeroes after the decimal point and before the 1). The physicists write 1.6×10^{-19} to represent this. It is the energy gained by a single electron in an electric field of 1 V. Machines called *particle accelerators* (or *atom smashers*) are rated by millions of electron volts (MeV), billions of electron volts (GeV), or trillions of electron volts (TeV) of energy capacity.

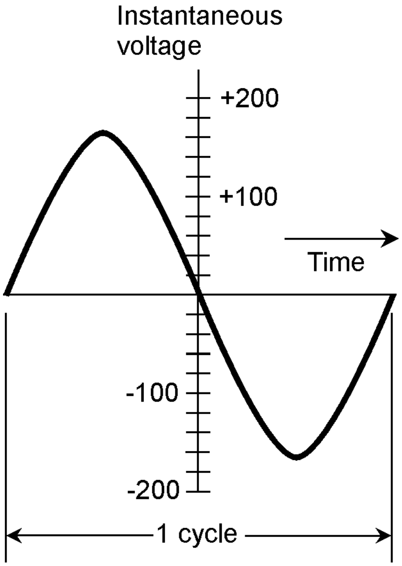
Another energy unit, employed to denote work, is the *foot-pound* (ft-lb). This is the work needed to raise a weight of one pound by a distance of one foot, not including any friction. It's equal to 1.356 joules.

All of these units, and conversion factors, are given in Table 2-3. Kilowatt-hours and watt-hours are also included in this table. In electricity and electronics, you need to be concerned only with the watt-hour and the kilowatt-hour for most purposes.

Alternating Current and the Hertz

This chapter, and this whole first section, is mostly concerned with *direct current* (dc). That's electric current that always flows in the same direction and that does not change in intensity (at least not too rapidly) with time. But household utility current is not of this kind. It reverses direction periodically, exactly once every 1/120 second. It goes through a complete cycle every 1/60 second. Every repetition is identical to every other. This is *alternating current* (ac).

Figure 2-8 shows the characteristic wave of ac, as a graph of voltage versus time. Notice that the maximum positive and negative voltages are not 117 V, as you've heard about household electricity, but close to 165 V. There is a reason for this difference. The *effective voltage* for an ac wave is never the same as the *instantaneous maximum*, or *peak*, voltage. In fact, for the common waveform shown in Fig. 2-8, the effective value is 0.707 times the peak value. Conversely, the peak value is 1.414 times the effective value.



2-8 One cycle of utility alternating current (ac). The instantaneous voltage is the voltage at any particular instant in time. The peak voltages are approximately plus and minus 165 V.

Because the whole cycle repeats itself every 1/60 second, the *frequency* of the utility ac wave is said to be 60 *hertz*, abbreviated 60 Hz. The German word *hertz* literally translates to "cycles per second." In the United States, this is the standard frequency for ac. In some places it is 50 Hz.

In wireless communications, higher frequencies are common, and you'll hear about *kilohertz* (kHz), *megahertz* (MHz), and *gigahertz* (GHz). The relationships among these units are as follows:

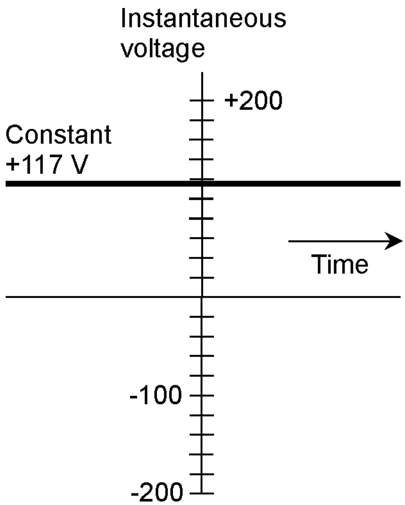
$$1 \text{ kHz} = 1000 \text{ Hz}$$
$$1 \text{ MHz} = 1000 \text{ kHz} = 1{,}000{,}000 \text{ Hz}$$
$$1 \text{ GHz} = 1000 \text{ MHz} = 1{,}000{,}000 \text{ kHz} = 1{,}000{,}000{,}000 \text{ Hz}$$

Usually, but not always, the waveshapes are of the type shown in Fig. 2-8. This waveform is known as a *sine wave* or a *sinusoidal* waveform.

Rectification and Pulsating Direct Current

Batteries and other sources of direct current (dc) produce constant voltage. This can be represented by a straight, horizontal line on a graph of voltage versus time (Fig. 2-9). For pure dc, the peak and effective values are identical. But sometimes the value of dc voltage fluctuates rapidly with time. This happens, for example, if the waveform in Fig. 2-8 is passed through a *rectifier* circuit.

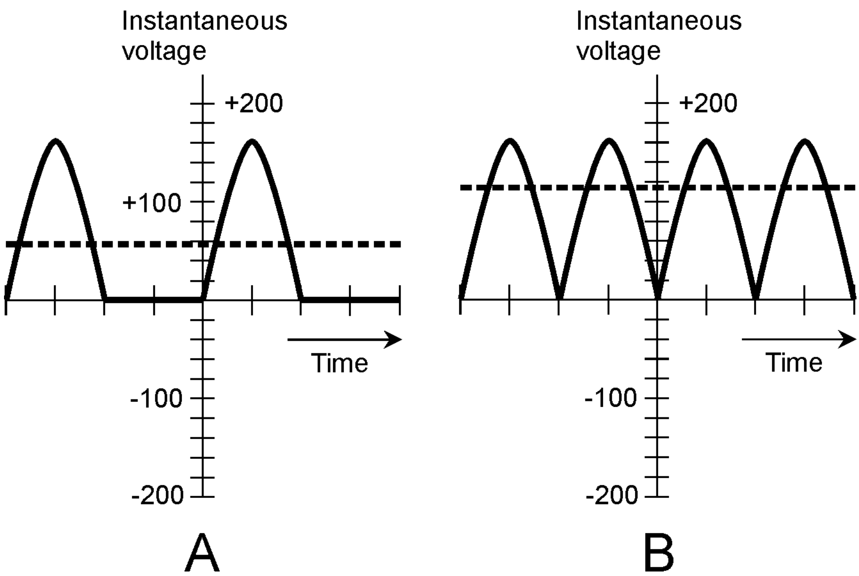
Rectification is a process in which ac is changed to dc. The most common method of doing this uses a device called the *diode.* Right now, you need not be concerned with how the rectifier circuit is put together. The point is that part of the ac wave is either cut off, or turned around upside down, so the output is *pulsating dc.* Figure 2-10 illustrates two different waveforms of pulsating dc. In the waveform at A, the negative (bottom) part has been cut off. At B, the negative portion of the wave has been inverted and made positive. The situation at A is known as *half-wave rectification*, because it involves only half the waveform. At B, the ac has been subjected to *full-wave rectification*, because



2-9 A representation of pure direct current (dc).

all of the original current still flows, even though the alternating nature has been changed so that the current never reverses.

The effective value, compared with the peak value, for pulsating dc depends on whether half-wave or full-wave rectification is applied to an ac wave. In Fig. 2-10A and B, effective voltage is shown as dashed lines, and the *instantaneous voltage* is shown as solid curves. The instantaneous voltage changes all the time, from instant to instant. (That's how it gets this name!) The peak voltage is



2-10 At A, half-wave rectification of common utility ac. At B, full-wave rectification of common utility ac. Effective voltages are shown by the dashed lines.

the maximum instantaneous voltage. Instantaneous voltage is never any greater than the peak voltage for any wave.

In Fig. 2-10B, the effective voltage is 0.707 times the peak voltage, just as is the case with ordinary ac. The direction of current flow, for many kinds of devices, doesn't make any difference. But in Fig. 2-10A, half of the wave has been lost. This cuts the effective value in half, so that it's only 0.354 times the peak value.

In household ac that appears in wall outlets for conventional appliances in the United States, the peak voltage is about 165 V; the effective value is 117 V. If full-wave rectification is used, the effective value is still 117 V. If half-wave rectification is used, the effective voltage is about 58.5 V.

Safety Considerations in Electrical Work

For our purposes, one rule applies concerning safety around electrical apparatus:

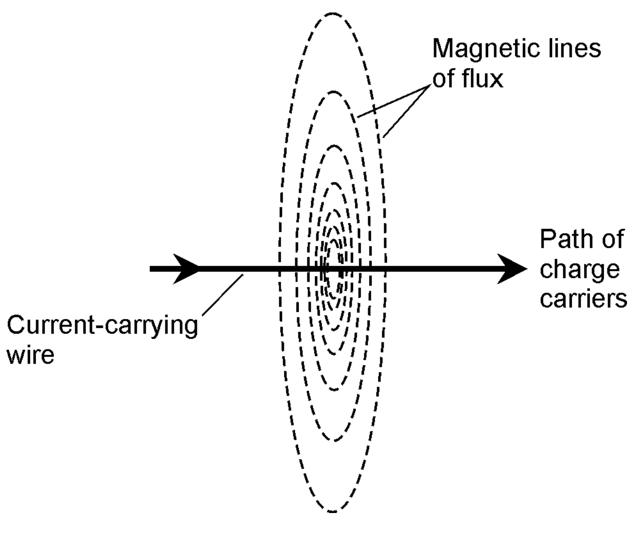
> *If you have any doubt about whether or not something is safe, leave it alone. Let a professional electrician work on it.*

Household voltage, normally about 117 V (but sometimes twice that for large appliances such as electric ranges and laundry machines), is more than sufficient to kill you if it appears across your chest cavity. Certain devices, such as automotive spark coils, can produce lethal currents even from the low voltage (12 to 14 V) in a car battery.

Consult the American Red Cross or your electrician concerning what kinds of circuits, procedures, and devices are safe and which aren't.

Magnetism

Electric currents and magnetic fields are closely related. Whenever an electric current flows—that is, when charge carriers move—a magnetic field accompanies the current. In a straight wire that carries electrical current, *magnetic lines of flux* surround the wire in circles, with the wire at the center, as shown in Fig. 2-11. (The lines of flux aren't physical objects; this is just a convenient way to rep-



2-11 Magnetic flux lines around a straight, current-carrying wire. The arrows indicate current flow.

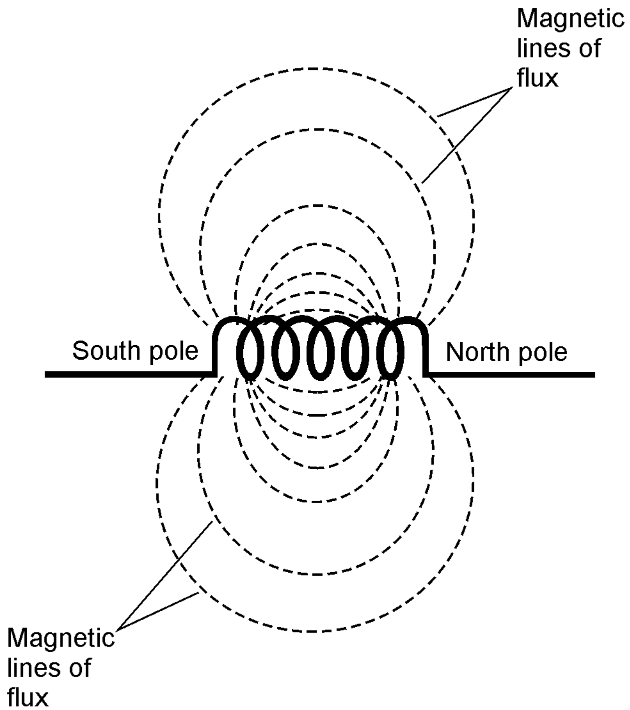
resent the magnetic field.) You'll sometimes hear or read about a certain number of flux lines per unit cross-sectional area, such as 100 lines per square centimeter. This is a relative way of talking about the intensity of the magnetic field.

Magnetic fields are produced when the atoms of certain materials align themselves. Iron is the most common metal that has this property. The atoms of iron in the core of the earth have become aligned to some extent; this is a complex interaction caused by the rotation of our planet and its motion with respect to the magnetic field of the sun. The magnetic field surrounding the earth is responsible for various effects, such as the concentration of charged particles that you see as the *aurora borealis* just after a solar eruption.

When a wire is coiled up, the resulting magnetic flux takes a shape similar to the flux field surrounding the earth, or the flux field around a bar magnet. Two well-defined *magnetic poles* develop, as shown in Fig. 2-12.

The intensity of a magnetic field can be greatly increased by placing a special core inside of a coil. The core should be of iron or some other material that can be readily magnetized. Such substances are called *ferromagnetic.* A core of this kind cannot actually increase the total quantity of magnetism in and around a coil, but it will cause the lines of flux to be much closer together inside the material. This is the principle by which an electromagnet works. It also makes possible the operation of electrical transformers for utility current.

Magnetic lines of flux are said to emerge from the magnetic north pole, and to run inward toward the magnetic south pole.



2-12 Magnetic flux lines around a current-carrying coil of wire. The flux lines converge at the magnetic poles.

Magnetic Units

The overall magnitude of a magnetic field is measured in units called *webers*, abbreviated Wb. One weber is mathematically equivalent to one volt-second (1 V · s). For weaker magnetic fields, a smaller unit, called the *maxwell* (Mx), is used. One maxwell is equal to 0.00000001 (one hundred-millionth) of a weber, or 0.01 microvolt-second (0.01 μV · s).

The *flux density* of a magnetic field is given in terms of webers or maxwells per square meter or per square centimeter. A flux density of one weber per square meter (1 Wb/m^2) is called one *tesla* (1 T). One *gauss* (1 G) is equal to 0.0001 T, or one maxwell per square centimeter (1 Mx/cm^2).

In general, as the electric current through a wire increases, so does the flux density near the wire. A coiled wire produces a greater flux density for a given current than a single, straight wire. And the more turns in the coil, the stronger the magnetic field will be.

Sometimes, magnetic field strength is specified in terms of *ampere-turns* (At). This is actually a unit of *magnetomotive force.* A one-turn wire loop, carrying 1 A of current, produces a field of 1 At. Doubling the number of turns, or the current, doubles the number of ampere-turns. Therefore, if you have 10 A flowing in a 10-turn coil, the magnetomotive force is 10 × 10, or 100 At. Or, if you have 100 mA flowing in a 100-turn coil, the magnetomotive force is 0.1 × 100, or 10 At. (Remember that 100 mA = 0.1 A.)

A less common unit of magnetomotive force is the *gilbert* (Gb). This unit is the equivalent of 0.796 At. Conversely, 1 At = 1.26 Gb.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct answers. The answers are listed in the back of this book.

1. A positive electric pole
   (a) has a deficiency of electrons.
   (b) has fewer electrons than the negative pole.
   (c) has an excess of electrons.
   (d) has more electrons than the negative pole.

2. An EMF of 1 V
   (a) cannot drive much current through a circuit.
   (b) represents a low resistance.
   (c) can sometimes produce a large current.
   (d) drops to zero in a short time.

3. A potentially lethal electric current is on the order of
   (a) 0.01 mA.
   (b) 0.1 mA.
   (c) 1 mA.
   (d) 0.1 A.

4. A current of 25 A is most likely drawn by
   (a) a flashlight bulb.
   (b) a typical household.
   (c) a utility power plant.
   (d) a small radio set.
5. A piece of wire has a conductance of 20 S. Its resistance is
   (a) 20 Ω.
   (b) 0.5 Ω.
   (c) 0.05 Ω.
   (d) 0.02 Ω.
6. A resistor has a value of 300 Ω. Its conductance is
   (a) 3.33 mS.
   (b) 33.3 mS.
   (c) 333 μS.
   (d) 0.333 S.
7. A span of wire 1 km long has a conductance of 0.6 S. What is the conductance of a span of this same wire that is 3 km long?
   (a) 1.8 S
   (b) 0.6 S
   (c) 0.2 S
   (d) More information is necessary to determine this.
8. Approximately how much current can a 2-kW generator reliably deliver at 117 V?
   (a) 17 mA
   (b) 234 mA
   (c) 17 A
   (d) 234 A
9. A circuit breaker is rated for 15 A at 117 V. Approximately how much power does this represent?
   (a) 1.76 kW
   (b) 1760 kW
   (c) 7.8 kW
   (d) 0.0078 kW
10. You are told that an air conditioner has cooled a room by 500 Btu over a certain period of time. What is this amount of energy in kWh?
    (a) 147 kWh
    (b) 14.7 kWh
    (c) 1.47 kWh
    (d) 0.147 kWh

11. Of the following energy units, the one most often used to define electrical energy is
    (a) the Btu.
    (b) the erg.
    (c) the foot-pound.
    (d) the kilowatt-hour.
12. The frequency of common household ac in the United States is
    (a) 60 Hz.
    (b) 120 Hz.
    (c) 50 Hz.
    (d) 100 Hz.
13. Half-wave rectification means that
    (a) half of the ac wave is inverted.
    (b) half of the ac wave is cut off.
    (c) the whole ac wave is inverted.
    (d) the effective voltage is half the peak voltage.
14. In the output of a half-wave rectifier,
    (a) half of the ac input wave is inverted.
    (b) the effective voltage is less than that of the ac input wave.
    (c) the effective voltage is the same as that of the ac input wave.
    (d) the effective voltage is more than that of the ac input wave.
15. In the output of a full-wave rectifier,
    (a) half of the ac input wave is inverted.
    (b) the effective voltage is less than that of the ac input wave.
    (c) the effective voltage is the same as that of the ac input wave.
    (d) the effective voltage is more than that of the ac input wave.
16. A low voltage, such as 12 V,
    (a) is never dangerous.
    (b) is always dangerous.
    (c) is dangerous if it is ac, but not if it is dc.
    (d) can be dangerous under certain conditions.
17. Which of the following units can represent magnetomotive force?
    (a) The volt-turn
    (b) The ampere-turn
    (c) The gauss
    (d) The gauss-turn

18. Which of the following units can represent magnetic flux density?
    (a) The volt-turn
    (b) The ampere-turn
    (c) The gauss
    (d) The gauss-turn
19. A ferromagnetic material
    (a) concentrates magnetic flux lines within itself.
    (b) increases the total magnetomotive force around a current-carrying wire.
    (c) causes an increase in the current in a wire.
    (d) increases the number of ampere-turns in a wire.
20. A coil has 500 turns and carries 75 mA of current. The magnetomotive force is
    (a) 37,500 At.
    (b) 375 At.
    (c) 37.5 At.
    (d) 3.75 At.

3
CHAPTER

Measuring Devices

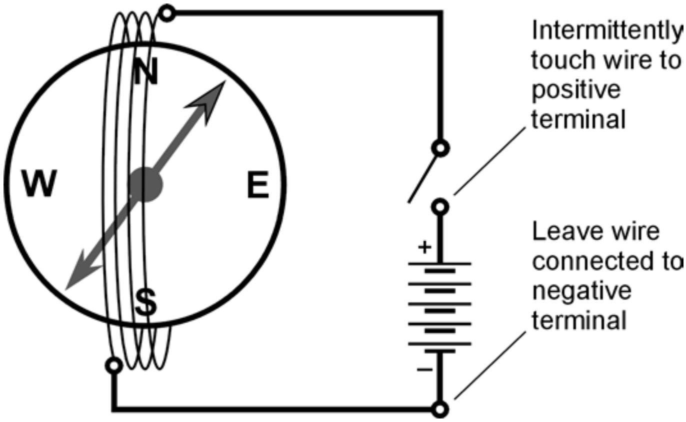
NOW THAT YOU'RE FAMILIAR WITH THE PRIMARY UNITS COMMON IN ELECTRICITY AND ELECTRONICS, let's look at the instruments that are employed to measure these quantities. Many measuring devices work because electric and magnetic fields produce forces proportional to the intensity of the field. Such meters work by means of *electromagnetic deflection* or *electrostatic deflection.* Sometimes, electric current is measured by the extent of heat it produces in a resistance. Such meters work by *thermal heating* principles. Some meters have small motors whose speed depends on the measured quantity. The rotation rate, or the number of rotations in a given time, can be measured or counted. Still other kinds of meters tally up electronic pulses, sometimes in thousands, millions, or billions. These are *electronic counters.*

Electromagnetic Deflection

Early experimenters with electricity and magnetism noticed that an electric current produces a magnetic field. When a magnetic compass is placed near a wire carrying a direct electric current, the compass doesn't point toward magnetic north. The needle is displaced. The extent of the displacement depends on how close the compass is brought to the wire, and also on how much current the wire is carrying.

When this effect was first observed, scientists tried different arrangements to see how much the compass needle could be displaced, and how small a current could be detected. An attempt was made to obtain the greatest possible current-detecting sensitivity. Wrapping the wire in a coil around the compass resulted in a device that could indicate a tiny electric current (Fig. 3-1). This effect is known as *galvanism*, and the meter so devised was called a *galvanometer.* Once this device was made, the scientists saw that the extent of the needle displacement increased with increasing current. Then, the only challenge was to calibrate the galvanometer somehow, and to find a standard so a universal meter could be engineered.

You can make your own galvanometer. Buy a cheap compass, about 2 feet of insulated bell wire, and a 6-volt lantern battery. Set it up as shown in Fig. 3-1. Wrap the wire around the compass four or five times, and align the compass so that the needle points along the wire turns while the wire is disconnected from the battery. Connect one end of the wire to the negative (–) terminal of the bat-

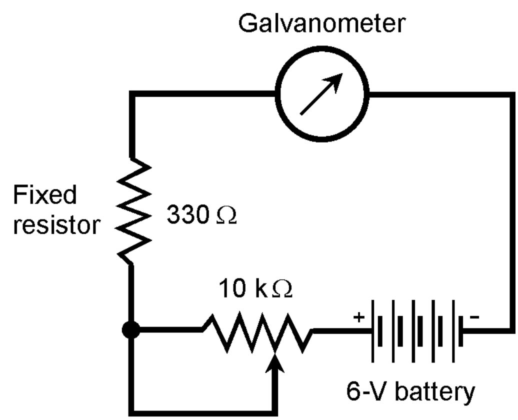


3-1 A simple galvanometer. The compass must lie flat.

tery. Touch the other end to the positive (+) terminal for a second or two, and watch the compass needle. Don't leave the wire connected to the battery for any length of time unless you want to drain the battery in a hurry.

You can buy a *resistor* and a *potentiometer* at a place like RadioShack, and set up an experiment that shows how galvanometers measure current. For a 6-V lantern battery, the fixed resistor should have a value of at least 330 Ω and should be rated for at least ¼ W. The potentiometer should have a maximum value of 10 kΩ. Connect the resistor and potentiometer in series between one end of the bell wire and one terminal of the battery, as shown in Fig. 3-2. The center contact of the potentiometer should be short-circuited to one of the end contacts, and the resulting two terminals used in the circuit.

When you adjust the potentiometer, the compass needle should deflect more or less, depending on the current through the wire. Early experimenters calibrated their meters by referring to the degrees scale around the perimeter of the compass.



3-2 A circuit for demonstrating how a galvanometer indicates relative current.

Electrostatic Deflection

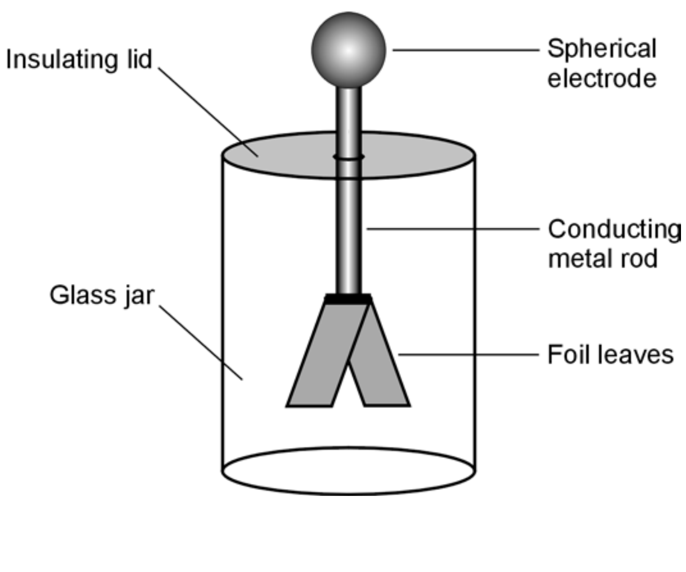
Electric fields produce forces, just as magnetic fields do. You have noticed this when your hair feels like it's standing on end in very dry or cold weather. You've heard that people's hair really does stand straight out just before a lightning bolt hits nearby. (This is no myth!)

The most common device for demonstrating electrostatic forces is the *electroscope.* It consists of two foil leaves, attached to a conducting rod, and placed in a sealed container so that air currents cannot move the foil leaves (Fig. 3-3). When a charged object is brought near, or touched to, the contact at the top of the rod, the leaves stand apart from each other. This is because the two leaves become charged with like electric poles—either an excess or a deficiency of electrons—and like poles always repel. The extent to which the leaves stand apart depends on the amount of electric charge. It is difficult to measure this deflection and correlate it with charge quantity; electroscopes do not make very good meters. But variations on this theme can be employed, so that electrostatic forces can operate against tension springs or magnets, and in this way, *electrostatic meters* can be made.

An electrostatic meter can quantify alternating (or ac) electric charges as well as direct (or dc) charges. This gives electrostatic meters an advantage over electromagnetic meters such as the galvanometers. If you connect a source of ac to the coil of the galvanometer device in Fig. 3-1 or Fig. 3-2, the compass needle will not give a clear deflection; current in one direction pulls the meter needle one way, and current in the other direction pushes the needle the opposite way. But if a source of ac is connected to an electrostatic meter, the plates repel whether the charge is positive or negative at any given instant in time.

Most electroscopes aren't sensitive enough to show much deflection with ordinary 117-V utility ac. Don't try connecting 117 V to an electroscope anyway. It can present an electrocution hazard if you bring it out to points where you can easily come into physical contact with it.

An electrostatic meter has another property that is sometimes an advantage in electrical or electronic work. This is the fact that the device does not draw any current, except a tiny initial current needed to put a charge on the plates. Sometimes, an engineer or experimenter doesn't want a measuring device to draw current, because this affects the behavior of the circuit under test. Galvanometers, by contrast, always need some current to produce an indication.



3-3 An electroscope can detect the presence of an electrostatic charge.

If you have access to a laboratory electroscope, try charging it up with a glass rod that has been rubbed against a cloth. When the rod is pulled away from the electroscope, the foil leaves remain standing apart. The charge just sits there! If the electroscope drew any current, the leaves would fall back together again, just as the galvanometer compass needle returns to magnetic north the instant you take the wire from the battery.

Thermal Heating

Another phenomenon, sometimes useful in the measurement of electric currents, is the fact that whenever current flows through a conductor having any resistance, that conductor is heated. All conductors have some resistance; none are perfect. The extent of this heating is proportional to the amount of current being carried by the wire.

By choosing the right metal or alloy, and by making the wire a certain length and diameter, and by employing a sensitive thermometer, and by putting the entire assembly inside a thermally insulating package, a *hot-wire meter* can be made. The hot-wire meter can measure ac as well as dc, because the current-heating phenomenon does not depend on the direction of current flow.

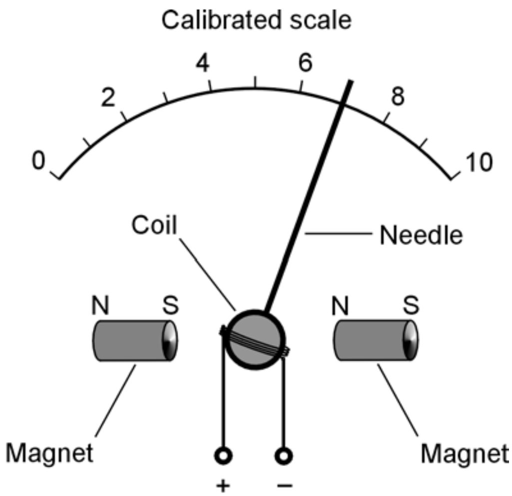
A variation of the hot-wire principle can be used to advantage by placing two different metals into contact with each other. If the right metals are chosen, the junction heats up when a current flows through it. This is called the *thermocouple principle.* As with the hot-wire meter, a thermometer can be used to measure the extent of the heating. But there is also another effect. A thermocouple, when it gets warm, generates dc. This dc can be measured with a galvanometer. This method is useful when it is necessary to have a fast meter *response time.*

The hot-wire and thermocouple effects are sometimes used to measure ac at high frequencies, in the range of hundreds of kilohertz up to tens of gigahertz.

Ammeters

A magnetic compass doesn't make a very convenient meter. It has to be lying flat, and the coil has to be aligned with the compass needle when there is no current. But of course, electrical and electronic devices aren't all oriented so as to be aligned with the north geomagnetic pole! But the external magnetic field doesn't have to come from the earth. It can be provided by a permanent magnet near or inside the meter. This supplies a stronger magnetic force than does the earth's magnetic field, and therefore makes it possible to make a meter that can detect much weaker currents. Such a meter can be turned in any direction, and its operation is not affected. The coil can be attached directly to the meter pointer, and suspended by means of a spring in the field of the magnet. This type of metering scheme, called the *D'Arsonval movement*, has been around since the earliest days of electricity, but it is still used in some metering devices today. The assembly is shown in Fig. 3-4. This is the basic principle of the *ammeter.*

A variation of the D'Arsonval movement can be obtained by attaching the meter needle to a permanent magnet, and winding the coil in a fixed form around the magnet. Current in the coil produces a magnetic field, and this in turn generates a force if the coil and magnet are aligned correctly with respect to each other. This works all right, but the mass of the permanent magnet causes a slower needle response. This type of meter is also more prone to *overshoot* than the true D'Arsonval movement; the inertia of the magnet's mass, once overcome by the magnetic force, causes the needle to fly past the actual point for the current reading, and then to wag back and forth a couple of times before coming to rest in the right place.

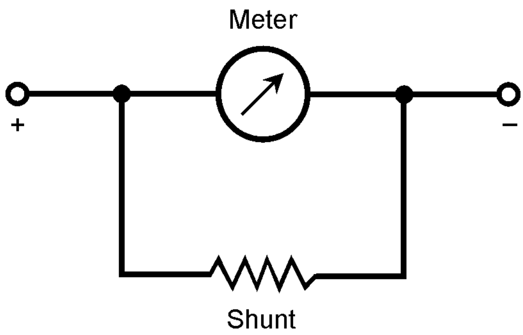


3-4 A functional drawing of a D'Arsonval meter movement (spring bearing not shown).

It is possible to use an *electromagnet* in place of the permanent magnet in the meter assembly. This electromagnet can be operated by the same current that flows in the coil attached to the meter needle. This gets rid of the need for a massive, permanent magnet inside the meter. It also eliminates the possibility that the meter sensitivity will change in case the strength of the permanent magnet deteriorates (such as might be caused by heat, or by severe mechanical vibration). The electromagnet can be either in series with, or in parallel with, the meter movement coil.

The sensitivity of the D'Arsonval-type meter, and of similar designs, depends on several factors. First is the strength of the permanent magnet (if the meter uses a permanent magnet). Second is the number of turns in the coil. The stronger the magnet, and the larger the number of turns in the coil, the less current is needed in order to produce a given magnetic force. If the meter is of the electromagnet type, the combined number of coil turns affects the sensitivity. Remember that the strength of a magnetomotive force is given in terms of ampere-turns. For a given current (number of amperes), the force increases in direct proportion to the number of coil turns. The more force in a meter, the greater the needle deflection for a given amount of current, and the smaller the current necessary to cause a certain amount of needle movement. The most sensitive ammeters can detect currents of just a microampere or two. The amount of current for *full-scale deflection* (the needle goes all the way up without banging against the stop pin) can be as little as about 50 μA in commonly available meters.

Sometimes, it is desirable to have an ammeter that will allow for a wide range of current measurements. The full-scale deflection of a meter assembly cannot easily be changed, because that would mean changing the number of coil turns and/or the strength of the magnet. But all ammeters have a certain amount of *internal resistance.* If a resistor, having the same internal resistance as the meter, is connected in parallel with the meter, the resistor will draw half the current. Then it will take twice the current through the assembly to deflect the meter to full scale, as compared with the meter alone. By choosing a resistor of just the right value, the full-scale deflection of an ammeter can be increased by a large factor, such as 10, or 100, or 1000. This resistor must be capable of carrying the current without burning up. It might have to draw practically all of the current flowing through the assembly, leaving the meter to carry only 1/10, or 1/100, or 1/1000 of the current. This is called a *shunt resistance* or *meter shunt* (Fig. 3-5). Meter shunts are used when it is necessary to measure very large currents, such as hundreds of amperes. They also allow microammeters or milliammeters to be used in a versatile *multimeter*, with many current ranges.



3-5 A resistor, called a *meter shunt*, can be connected across a current-detecting meter to reduce the sensitivity.

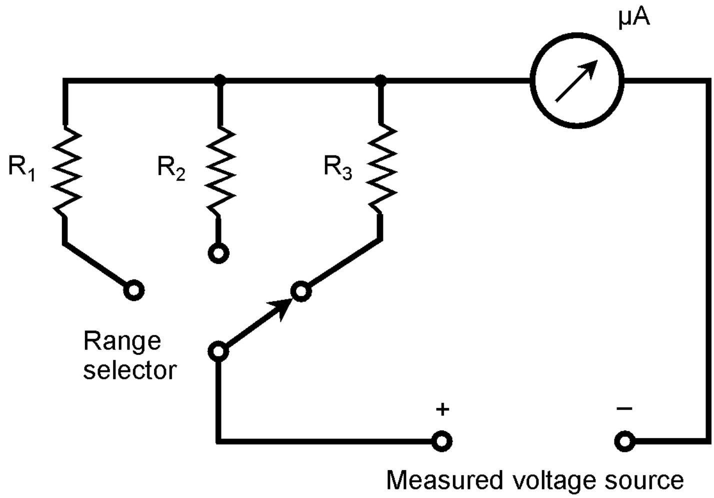
Voltmeters

Current, as we have seen, consists of a flow of charge carriers. Voltage, or electromotive force (EMF), or potential difference, is the "pressure" that makes current possible. Given a circuit whose resistance is constant, the current that flows in the circuit is directly proportional to the voltage placed across it. Early electrical experimenters recognized that an ammeter could be used to measure voltage, because an ammeter is a form of constant-resistance circuit. If you connect an ammeter directly across a source of voltage such as a battery, the meter needle deflects. In fact, a milliammeter needle will probably be "pinned" if you do this with it, and a microammeter might well be wrecked by the force of the needle striking the pin at the top of the scale. For this reason, you should never connect milliammeters or microammeters directly across voltage sources. An ammeter, perhaps with a range of 0 to 10 A, might not deflect to full scale if it is placed across a battery, but it's still a bad idea to do this, because it will rapidly drain the battery. Some batteries, such as automotive lead-acid cells, can explode under these conditions.

Ammeters have low internal resistance. They are designed that way deliberately. They are meant to be connected in series with other parts of a circuit, not right across a power supply. But if you place a large resistor in series with an ammeter, and then connect the ammeter across a battery or other type of power supply, you no longer have a short circuit. The ammeter will give an indication that is directly proportional to the voltage of the supply. The smaller the full-scale reading of the ammeter, the larger the resistance that is needed to get a meaningful indication on the meter. Using a microammeter and a very large value of resistance in series, a *voltmeter* can be devised that will draw only a little current from the source.

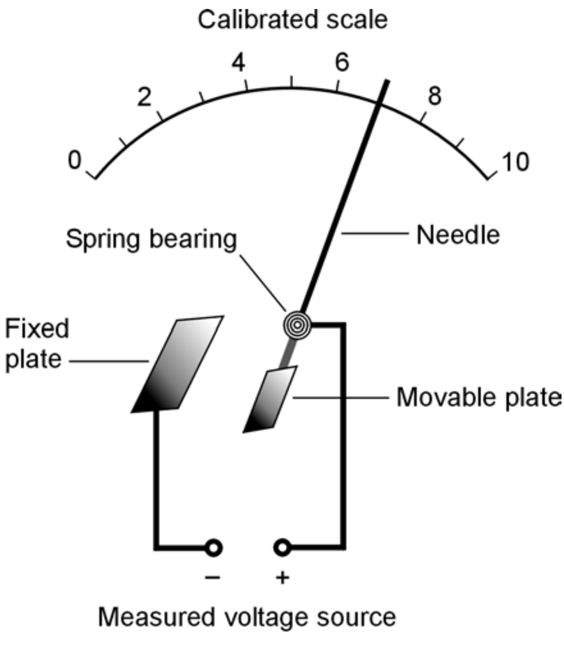
A voltmeter can be made to have various ranges for the full-scale reading, by switching different values of resistance in series with the microammeter (Fig. 3-6). The internal resistance of the meter is large because the values of the resistors are large. The greater the supply voltage, the larger the internal resistance of the meter, because the necessary series resistance increases as the voltage increases.

A voltmeter should have high internal resistance, and the higher the better! The reason for this is that you don't want the meter to draw much current from the power source. This current should go, as much as possible, toward operating whatever circuit is hooked up to the power supply, and not into getting a reading of the voltage. Also, you might not want, or need, to have the voltmeter constantly connected in the circuit; you might need the voltmeter for testing many different circuits. You don't want the behavior of a circuit to be affected the instant you connect the voltmeter to the supply. The less current a voltmeter draws, the less it affects the behavior of anything that is working from the power supply.



3-6 A simple circuit using a microammeter (μA) to measure dc voltage.

A completely different type of voltmeter uses the effect of electrostatic deflection, rather than electromagnetic deflection. Remember that electric fields produce forces, just as do magnetic fields. Therefore, a pair of plates attract or repel each other if they are charged. The *electrostatic voltmeter* takes advantage of the attractive force between two plates having opposite electric charge, or having a large potential difference. Figure 3-7 is a simplified drawing of the mechanics of an electrostatic voltmeter. It draws almost no current from the power supply. The only thing between the plates is air, and air is a nearly perfect insulator. The electrostatic meter can indicate ac voltage as well as dc voltage. The construction tends to be fragile, however, and mechanical vibration can influence the reading.



3-7 A functional drawing of an electrostatic voltmeter movement.

Ohmmeters

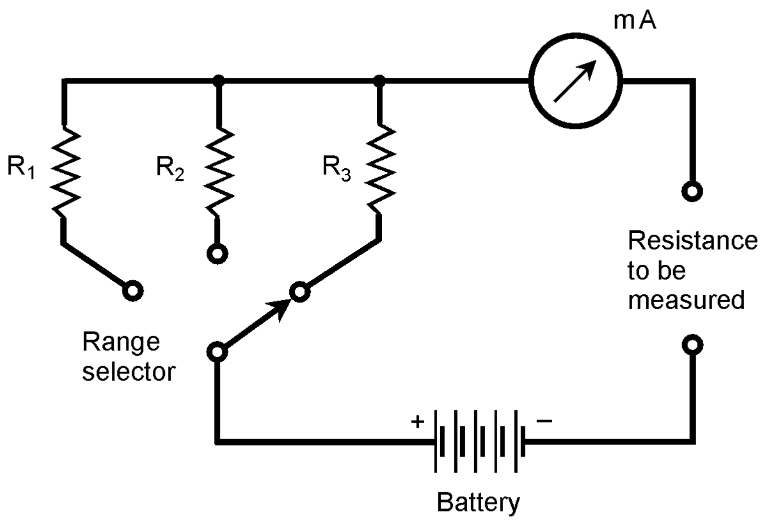
If all other factors are held constant, the current through a circuit depends on the resistance. This provides us with a means for measuring resistance. An *ohmmeter* can be constructed by placing a milliammeter or microammeter in series with a set of fixed, switchable resistances and a battery that provides a known, constant voltage (Fig. 3-8). By selecting the resistances appropriately, the meter gives indications in ohms over any desired range. The zero point on the milliammeter or microammeter is assigned the value of *infinity ohms*, meaning a perfect insulator. The full-scale value is set at a certain minimum, such as 1 Ω, 100 Ω, 1 kΩ, or 10 kΩ.

An ohmmeter must be calibrated at the factory where it is made, or in an electronics lab. A slight error in the values of the series resistors can cause gigantic errors in measured resistance. Therefore, precise *tolerances* are needed for these resistors. That means their values must actually be what the manufacturer claims they are, to within a fraction of 1 percent if possible. It is also necessary that the battery provide exactly the right voltage.

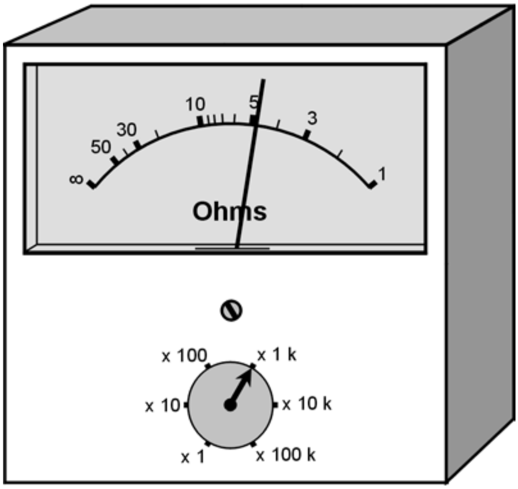
The scale of an ohmmeter is *nonlinear.* That means the graduations are not of the same width everywhere on the meter scale. The graduations tend to be squashed together toward the infinity end of the scale. Because of this, it is difficult to interpolate for high values of resistance unless the appropriate meter range is selected.

Engineers and technicians usually connect an ohmmeter in a circuit with the meter set for the highest resistance range first. Then they switch the range down until the meter needle is in a part of the scale that is easy to read. Finally, the reading is taken, and is multiplied (or divided) by the appropriate amount as indicated on the range switch. Figure 3-9 shows an ohmmeter reading. The meter itself indicates approximately 4.7, but the range switch says 1 kΩ. This indicates a resistance of about 4.7 kΩ, or 4700 Ω.

Ohmmeters give inaccurate readings if there is a voltage between the points where the meter is connected. This is because such a voltage either adds to, or subtracts from, the ohmmeter's own battery voltage. Sometimes, in this type of situation, an ohmmeter might tell you that a circuit has



3-8 A circuit using a milliammeter (mA) to measure dc resistance.



3-9 An example of an ohmmeter reading. This device shows about $4.7 \times 1\ \text{k}\Omega = 4.7\ \text{k}\Omega = 4700\ \Omega$.

"more than infinity" ohms! The needle will hit the pin at the left end of the scale. Therefore, when using an ohmmeter to measure resistance, you must always be sure that there is no voltage between the points under test. The best way to do this is to switch off the equipment in question.

Multimeters

In the electronics lab, a common piece of test equipment is the *multimeter*, in which different kinds of meters are combined into a single unit. The *volt-ohm-milliammeter* (VOM) is the most often used. As its name implies, it combines voltage, resistance, and current measuring capabilities. You should not have trouble envisioning how a single milliammeter can be used for measuring voltage, current, and resistance. The preceding discussions for measurements of these quantities have all included methods in which a current meter can be used to measure the intended quantity.

Commercially available multimeters have certain limits in the values they can measure. The maximum voltage is around 1000 V. The measurement of larger voltages requires special probes and heavily insulated wires, as well as other safety precautions. The maximum current that a common VOM can measure is about 1 A. The maximum measurable resistance is on the order of several megohms or tens of megohms. The lower limit of resistance indication is around 0.1 to 1 Ω.

FET Voltmeters

A good voltmeter disturbs the circuit under test as little as possible, and this requires that the meter have high internal resistance. Besides the electrostatic-type voltmeter, there is another way to get high internal resistance. This is to sample a tiny current, far too small for any meter to directly indicate, and then amplify this current so a conventional milliammeter or microammeter can display it. When a minuscule current is drawn from a circuit, the equivalent resistance is always extremely high.

The most effective way to accomplish voltage amplification, while making sure that the current drawn is exceedingly small, is to use a *field-effect transistor*, or FET. (Don't worry about how such amplifiers work right now; you'll learn all about that later in this book.) A voltmeter that uses a FET voltage amplifier to minimize the current drain is known as a *FET voltmeter* (FETVM). It has extremely high input resistance, along with good sensitivity and amplification.

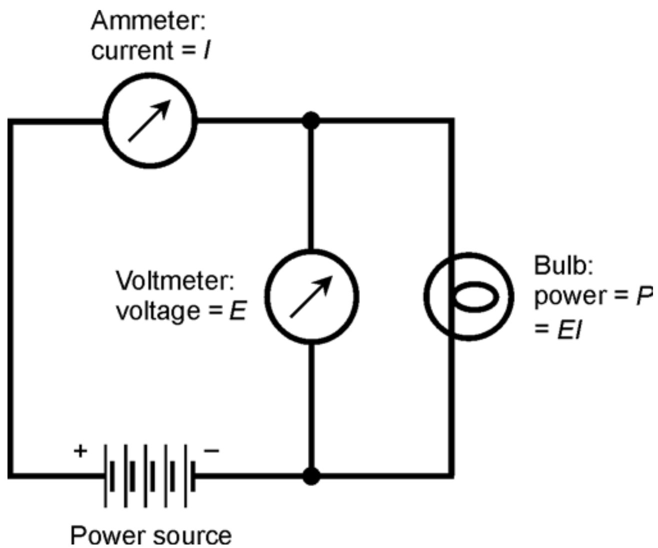
Wattmeters

The measurement of electrical power requires that voltage and current both be measured simultaneously. Remember that in a dc circuit, the power (P) in watts is the product of the voltage (E) in volts and the current (I) in amperes. That is, $P = EI$. In fact, watts are sometimes called *volt-amperes* in dc circuits.

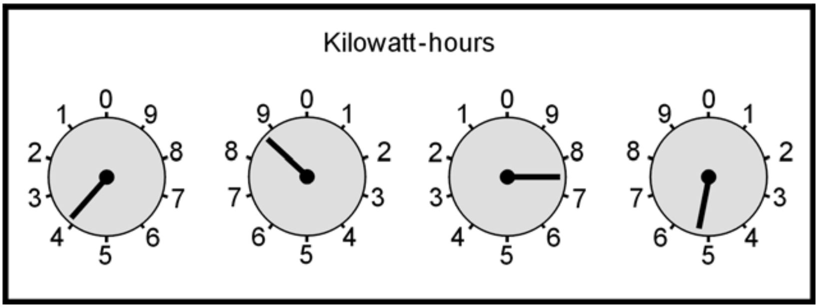
Do you think you can connect a voltmeter in parallel with a circuit, thereby getting a reading of the voltage across it, and also hook up an ammeter in series to get a reading of the current through the circuit, and then multiply volts times amperes to get watts consumed by the circuit? Well, you can. For most dc circuits, this is an excellent way to measure power, as shown in Fig. 3-10.

Sometimes, it's simpler yet. In many cases, the voltage from the power supply is constant and predictable. Utility power is a good example. The effective voltage is always very close to 117 V. Although it's ac, and not dc, power in most utility circuits can be measured in the same way as power is measured in dc circuits: by means of an ammeter connected in series with the circuit, and calibrated so that the multiplication (times 117) has already been done. Then, rather than 1 A, the meter will show a reading of 117 W, because $P = EI = 117 \times 1 = 117$ W. If the meter reading is 300 W, the current is $I = P/E = 300/117 = 2.56$ A. An electric iron might consume 1000 W, or a current of 1000/117 = 8.55 A. A large heating unit might gobble up 2000 W, requiring a current of 2000/117 = 17.1 A. You should not be surprised if this blows a fuse or trips a circuit breaker, because these devices are often rated for 15 A.

Specialized wattmeters are necessary for the measurement of radio-frequency (RF) power, or for peak audio power in a high-fidelity amplifier, or for certain other specialized applications. But almost all of these meters, whatever the associated circuitry, use simple ammeters, milliammeters, or microammeters as their indicating devices.



3-10 In a dc circuit, power can be measured with a voltmeter and an ammeter, connected as shown here.



3-11 A utility meter with four rotary analog dials. In this example, the reading is a little more than 3875 kWh.

Watt-Hour Meters

Electrical energy, as you now know, is measured in watt-hours or kilowatt-hours (kWh). Not surprisingly, a metering device that indicates energy in these units is called a *watt-hour meter* or a *kilowatt-hour meter.*

The most often used means of measuring electrical energy is by using a small electric motor, the speed of which depends on the current, and thereby on the power at a constant voltage. The number of turns of the motor shaft, in a given length of time, is directly proportional to the number of kilowatt-hours consumed. The motor is placed at the point where the utility wires enter the building. This is usually at a point where the voltage is 234 V. At this point the circuit is split into some circuits with 234 V (for heavy-duty appliances such as the oven, washer, and dryer) and general household circuits at 117 V (for smaller appliances such as lamps, clock radios, and television sets).

If you've observed a kilowatt-hour meter, you have seen a disk spinning, sometimes fast, other times slowly. Its speed depends on the power being used at any given time. The total number of turns of this little disk, every month, determines the size of the bill you will get, as a function also, of course, of the cost per kilowatt-hour.

Kilowatt-hour meters count the number of disk turns by means of geared rotary drums or pointers. The drum-type meter gives a direct digital readout. The pointer type has several scales calibrated from 0 to 9 in circles, some going clockwise and others going counterclockwise. Reading a pointer-type utility meter is a little tricky, because you must think in whatever direction (clockwise or counterclockwise) the scale goes. An example of a pointer-type utility meter is illustrated in Fig. 3-11. Read from left to right. For each meter scale, take down the number that the pointer has most recently passed. Write down the rest as you go. The meter shown in the figure reads a little more than 3875 kWh.

Digital Readout Meters

Increasingly, metering devices are being designed so that they provide a direct readout. The number on the meter is the indication. It's that simple. Such a meter is called a *digital meter.*

The main advantage of a digital meter is the fact that it's easy for anybody to read, and there is no chance for interpolation errors. This is ideal for utility meters, clocks, and some kinds of ammeters, voltmeters, and wattmeters. It works well when the value of the quantity does not change often or fast.

There are some situations in which a digital meter is a disadvantage. One good example is the signal-strength indicator in a radio receiver. This meter bounces up and down as signals fade, or as you tune the radio, or sometimes even as the signal modulates. A digital meter will show nothing but a constantly changing, meaningless set of numerals. Digital meters require a certain length of time to lock in to the current, voltage, power, or other quantity being measured. If this quantity never settles at any one value for a long enough time, the meter can never lock in.

Meters with a scale and pointer are known as *analog meters*. Their main advantages are that they allow interpolation, they give the operator a sense of the quantity relative to other possible values, and they follow along when a quantity changes. Some engineers and technicians prefer analog metering, even in situations where digital meters would work just as well.

One potential hang-up with digital meters is being certain of where the decimal point goes. If you're off by one decimal place, the error will be by a factor of 10. Also, you need to be sure you know what the units are. For example, a frequency indicator might be reading out in megahertz, and you might forget and think it is giving you a reading in kilohertz. That's a mistake by a factor of 1000! Of course, this latter type of error can happen with analog meters, too.

Frequency Counters

The measurement of energy used by your home is an application to which digital metering is well suited. A digital kilowatt-hour meter is easier to read than the pointer-type meter. When measuring frequencies of radio signals, digital metering is not only more convenient, but far more accurate.

A *frequency counter* measures the frequency of an ac wave by actually counting pulses, in a manner similar to the way the utility meter counts the number of turns of a motor. But the frequency counter works electronically, without any moving parts. It can keep track of thousands, millions, or billions of pulses per second, and it shows the rate on a digital display that is as easy to read as a digital watch.

The accuracy of the frequency counter is a function of the *lock-in time*. Lock-in is usually done in 0.1 second, 1 second, or 10 seconds. Increasing the lock-in time by a factor of 10 will cause the accuracy to increase by one additional digit. Modern frequency counters are good to six, seven, or eight digits; sophisticated lab devices can show frequency to nine or ten digits.

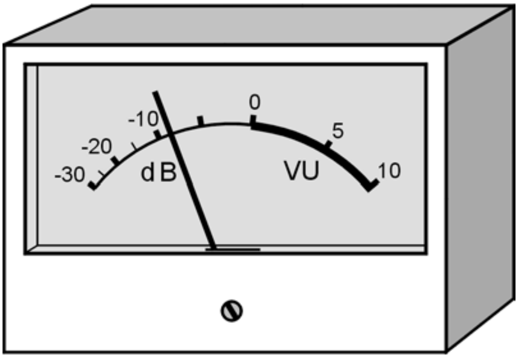
Other Meter Types

Here are a few of the less common types of meters that you will occasionally encounter in electrical and electronics applications.

VU and Decibel Meters

In high-fidelity equipment, especially the more sophisticated amplifiers ("amps"), *loudness meters* are sometimes used. These are calibrated in *decibels*, a unit that you will often have to use, and interpret, in reference to electronic signal levels. A decibel is an increase or decrease in sound or signal level that you can just barely detect, if you are expecting the change.

Audio loudness is given in *volume units* (VU), and the meter that indicates it is called a *VU meter*. The typical VU meter has a zero marker with a red line to the right and a black line to the left, and is calibrated in decibels (dB) below the zero marker and volume units above it (Fig. 3-12). The meter might also be calibrated in *watts rms*, an expression for audio power. As music is played



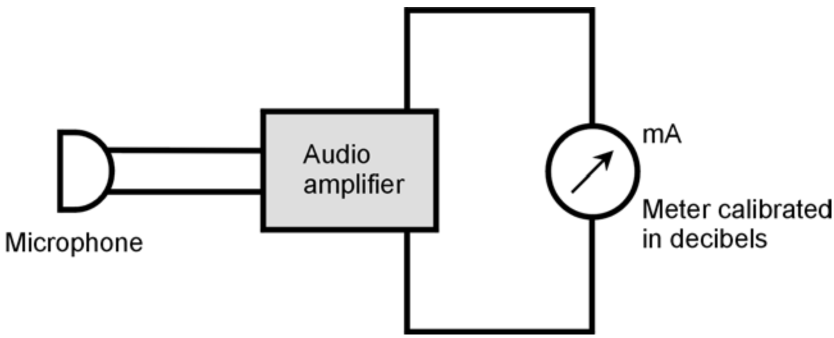
3-12 A VU (volume-unit) meter. The heavy portion of the scale (to the right of 0) is usually red, indicating the risk of audio distortion.

through the system, or as a voice comes over it, the VU meter needle kicks up. The amplifier volume should be kept down so that the meter doesn't go past the zero mark and into the red range. If the meter does kick up into the red scale, it means that distortion is taking place within the amplifier circuit.

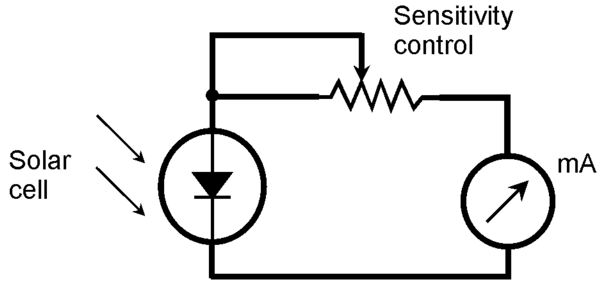
Sound level in general can be measured by means of a *sound-level meter*, calibrated in decibels (dB) and connected to the output of a precision amplifier with a microphone of known sensitivity (Fig. 3-13). Have you read that a vacuum cleaner will produce "80 dB" of sound, and a large truck going by will subject your ears to "90 dB"? These figures are determined by a sound-level meter, and are defined with respect to the *threshold of hearing*, which is the faintest sound that a person with good ears can hear.

Light Meters

The intensity of visible light is measured by means of a *light meter* or *illumination meter.* It is tempting to suppose that it's easy to make this kind of meter by connecting a milliammeter to a solar (photovoltaic) cell. As things work out, this is a good way to construct an inexpensive light meter (Fig. 3-14). More sophisticated devices use dc amplifiers, similar to the type found in a FETVM, to enhance sensitivity and to allow for several different ranges of readings.



3-13 A meter for measuring sound levels. The output of the audio amplifier is rectified to produce dc that the meter can detect.



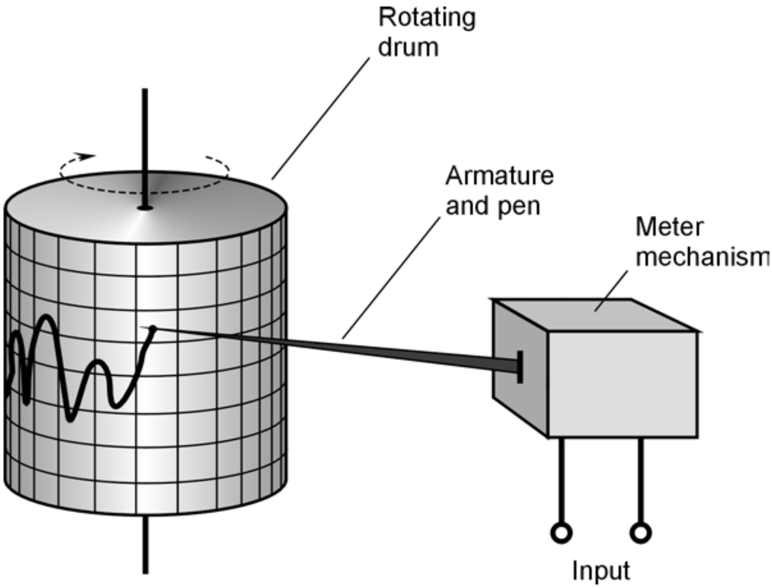
3-14 A simple light meter. A microammeter can be substituted for the milliammeter if greater sensitivity is required.

One problem with this design is that solar cells are not sensitive to light at exactly the same wavelengths as human eyes. This can be overcome by placing a colored filter in front of the solar cell, so that the solar cell becomes sensitive to the same wavelengths, in the same proportions, as human eyes. Another problem is calibrating the meter. This must usually be done at the factory, in standard illumination units such as *lumens* or *candela.*

With appropriate modification, meters such as the one in Fig. 3-14 can be used to measure *infrared* (IR) or *ultraviolet* (UV) intensity. Various specialized photovoltaic cells have peak sensitivity at nonvisible wavelengths, including IR and UV.

Pen Recorders

A meter movement can be equipped with a marking device to keep a graphic record of the level of some quantity with respect to time. Such a device is called a *pen recorder.* The paper, with a calibrated scale, is taped to a rotating drum. The drum, driven by a clock motor, turns at a slow rate, such as one revolution per hour or one revolution in 24 hours. A simplified drawing of a pen recorder is shown in Fig. 3-15.



3-15 A functional drawing of a pen recorder.

A device of this kind, along with a wattmeter, can be employed to get a reading of the power consumed by your household at various times during the day. In this way you can find out when you use the most power, and at what particular times you might be using too much.

Oscilloscopes

Another graphic metering device is the *oscilloscope*. This measures and records quantities that vary rapidly, at rates of hundreds, thousands, or millions of times per second. It creates a "graph" by throwing a beam of electrons at a phosphor screen. A *cathode-ray tube*, similar to the kind in a television set, is employed. Some oscilloscopes have electronic conversion circuits that allow for the use of a solid-state *liquid crystal display* (LCD).

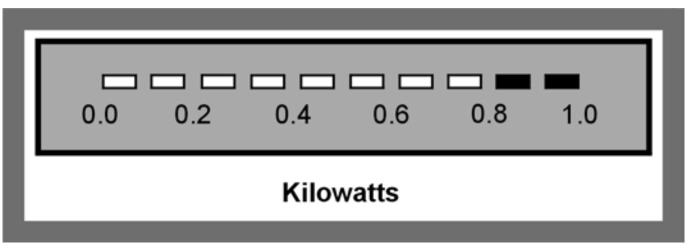
Oscilloscopes are useful for observing and analyzing the shapes of signal waveforms, and also for measuring peak signal levels (rather than just the effective levels). An oscilloscope can also be used to approximately measure the frequency of a waveform. The horizontal scale of an oscilloscope shows time, and the vertical scale shows the instantaneous signal voltage. An oscilloscope can indirectly measure power or current, by using a known value of resistance across the input terminals.

Technicians and engineers develop a sense of what a signal waveform should look like, and then they can often tell, by observing the oscilloscope display, whether or not the circuit under test is behaving the way it should. This is a subjective measurement, because it is qualitative as well as quantitative.

Bar-Graph Meters

A cheap, simple kind of meter can be made using a string of light-emitting diodes (LEDs) or an LCD along with a digital scale to indicate approximate levels of current, voltage, or power. This type of meter, like a digital meter, has no moving parts to break. To some extent, it offers the relative-reading feeling you get with an analog meter. Figure 3-16 is an example of a bar-graph meter that is used to show the power output, in kilowatts, for a radio transmitter. This meter can follow along quite well with rapid fluctuations in the reading. In this example, the meter indicates about 0.8 kW, or 800 W.

The chief drawback of the bar-graph meter is that it isn't very accurate. For this reason it is not generally used in laboratory testing. In addition, the LED or LCD devices sometimes flicker when the level is between two values given by the bars. This creates an illusion of circuit instability. With bright LEDs, it can also be quite distracting.



3-16 A bar-graph meter. In this case, the indication is about 80 percent of full-scale, representing 0.8 kW, or 800 W.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 out of 20 correct. Answers are in the back of the book.

1. The attraction or repulsion between two electrically charged objects is called
   (a) electromagnetic deflection.
   (b) electrostatic force.
   (c) magnetic force.
   (d) electroscopic force.

2. The change in the direction of a compass needle, when a current-carrying wire is brought near, is called
   (a) electromagnetic deflection.
   (b) electrostatic force.
   (c) magnetic force.
   (d) electroscopic force.

3. Suppose a certain current in a galvanometer causes the compass needle to deflect by 20 degrees, and then this current is doubled while the polarity stays the same. The angle of the needle deflection will
   (a) decrease.
   (b) stay the same.
   (c) increase.
   (d) reverse direction.

4. One important advantage of an electrostatic meter is the fact that
   (a) it measures very small currents.
   (b) it can handle large currents.
   (c) it can detect and indicate ac voltages as well as dc voltages.
   (d) it draws a large current from a power supply.

5. A thermocouple
   (a) gets warm when dc flows through it.
   (b) is a thin, straight, special wire.
   (c) generates dc when exposed to visible light.
   (d) generates ac when heated.

6. An important advantage of an electromagnet-type meter over a permanent-magnet meter is the fact that
   (a) the electromagnet meter costs much less.
   (b) the electromagnet meter need not be aligned with the earth's magnetic field.
   (c) the permanent-magnet meter has a more sluggish coil.
   (d) the electromagnet meter is more rugged.

7. Ammeter shunts are useful because
   (a) they increase meter sensitivity.
   (b) they make a meter more physically rugged.
   (c) they allow for measurement of large currents.
   (d) they prevent overheating of the meter movement.

8. Voltmeters should generally have
   (a) high internal resistance.
   (b) low internal resistance.
   (c) the greatest possible sensitivity.
   (d) the ability to withstand large currents.

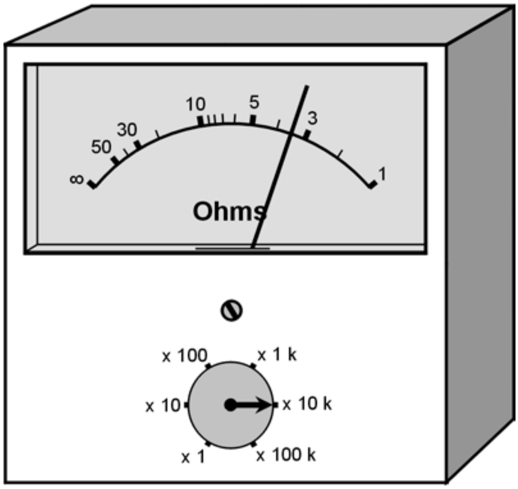
9. In order to measure the power-supply voltage that is applied to an electrical circuit, a voltmeter should be placed
   (a) in series with the circuit that works from the supply.
   (b) between the negative pole of the supply and the circuit working from the supply.
   (c) between the positive pole of the supply and the circuit working from the supply.
   (d) in parallel with the circuit that works from the supply.

10. Which of the following will *not* normally cause a *large* error in an ohmmeter reading?
   (a) A small voltage between points under test
   (b) A slight change in switchable internal resistance
   (c) A small change in the resistance to be measured
   (d) A slight error in the range switch position

11. The ohmmeter in Fig. 3-17 shows a reading of approximately
   (a) 34,000 Ω.
   (b) 3.4 kΩ.
   (c) 340 Ω.
   (d) 34 Ω.

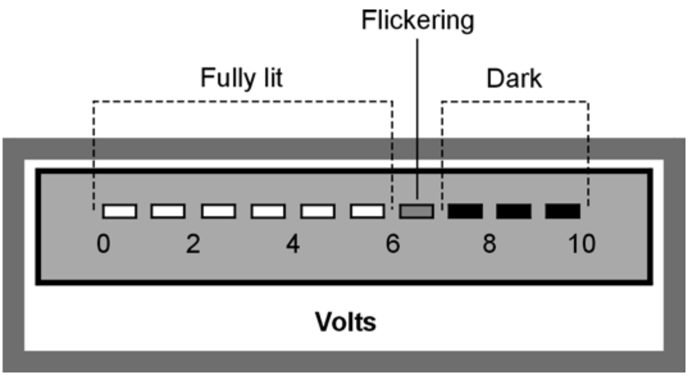
12. The main advantage of a FETVM over a conventional voltmeter is the fact that the FETVM
   (a) can measure lower voltages.
   (b) draws less current from the circuit under test.
   (c) can withstand higher voltages safely.
   (d) is sensitive to ac voltage as well as to dc voltage.

13. Which of the following is *not* a function of a fuse?
   (a) To ensure there is enough current available for an appliance to work right
   (b) To make it impossible to use appliances that are too large for a given circuit
   (c) To limit the amount of power that a device can draw from the electrical circuit
   (d) To make sure the current drawn by an appliance cannot exceed a certain limit



3-17 Illustration for Quiz Question 11.

14. A utility meter's motor speed depends directly on
    (a) the number of ampere-hours being used at the time.
    (b) the number of watt-hours being used at the time.
    (c) the number of watts being used at the time.
    (d) the number of kilowatt-hours being used at the time.
15. A utility meter's readout indicates
    (a) voltage.
    (b) power.
    (c) current.
    (d) energy.
16. A typical frequency counter
    (a) has an analog readout.
    (b) is accurate to six digits or more.
    (c) works by indirectly measuring current.
    (d) works by indirectly measuring voltage.
17. A VU meter is *never* used to get a general indication of
    (a) sound intensity.
    (b) decibels.
    (c) power in an audio amplifier.
    (d) visible light intensity.



3-18 Illustration for Quiz Question 20.

18. The meter movement in an illumination meter directly measures
    (a) current.
    (b) voltage.
    (c) power.
    (d) energy.
19. An oscilloscope *cannot* be used to indicate
    (a) frequency.
    (b) wave shape.
    (c) energy.
    (d) peak signal voltage.
20. What voltage would be expected to produce the reading on the bar-graph meter shown in Fig. 3-18?
    (a) 6.0 V
    (b) 6.5 V
    (c) 7.0 V
    (d) There is no way to tell because the meter, as shown, is malfunctioning.

4
CHAPTER

Direct-Current Circuit Basics

YOU'VE ALREADY SEEN SOME SIMPLE ELECTRICAL CIRCUIT DIAGRAMS. IN THIS CHAPTER, YOU'LL GET more acquainted with this type of diagram. You'll also learn more about how current, voltage, resistance, and power are related in dc and low-frequency ac circuits.

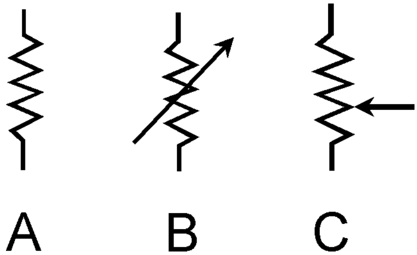
Schematic Symbols

In this course, the idea is to familiarize you with schematic symbols by getting you to read and use them in action. But right now, why not check out Appendix B, which is a comprehensive table of symbols? Then refer to it frequently in the future, especially when you see a symbol you don't remember or recognize.

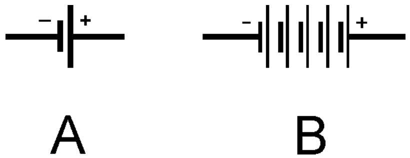
The simplest schematic symbol is the one representing a wire or *electrical conductor*: a straight, solid line. Sometimes, dashed lines are used to represent conductors, but usually, dashed lines are drawn to partition diagrams into constituent circuits, or to indicate that certain components interact with each other or operate in step with each other. Conductor lines are almost always drawn either horizontally across or vertically up and down the page. This keeps the diagram neat and easy to read.

When two conductor lines cross, they aren't connected at the crossing point unless a heavy black dot is placed where the two lines meet. The dot should always be clearly visible wherever conductors are to be connected, no matter how many of them meet at the junction. A *resistor* is indicated by a zigzag. A variable resistor, or *potentiometer*, is indicated by a zigzag with an arrow through it, or by a zigzag with an arrow pointing at it. These symbols are shown in Fig. 4-1.

4-1 Schematic symbols for a fixed resistor (A), a two-terminal variable resistor (B), and a three-terminal potentiometer (C).

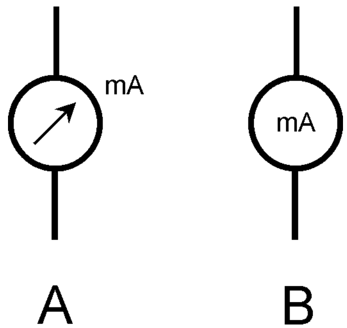


An *electrochemical cell* (such as a common dime-store battery) is shown by two parallel lines, one longer than the other. The longer line represents the plus terminal. A true *battery*, which is a combination of two or more cells in series, is indicated by several parallel lines, alternately long and short. It's not necessary to use more than four lines to represent a battery, although you'll often see 6, 8, 10, or even 12 lines. Symbols for a cell and a battery are shown in Fig. 4-2.



4-2 Schematic symbols for an electrochemical cell (A) and an electrochemical battery (B).

Meters are portrayed as circles. Sometimes the circle has an arrow inside it, and the meter type, such as mA (milliammeter) or V (voltmeter) is written alongside the circle, as shown in Fig. 4-3A. Sometimes the meter type is indicated inside the circle, and there is no arrow (Fig. 4-3B). It doesn't matter which way you draw them, as long as you're consistent throughout a schematic diagram.

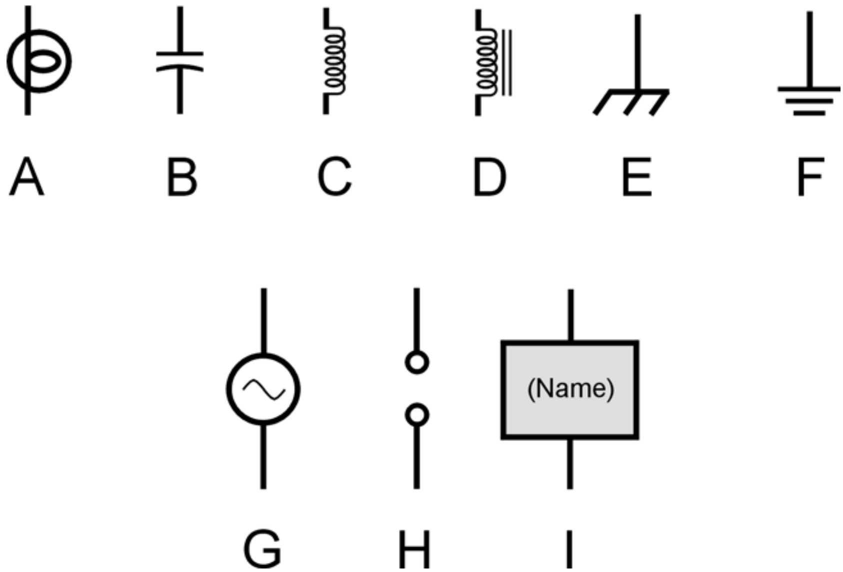


4-3 Meter symbols can have the designator either outside the circle (A) or inside (B). In this case, both symbols represent a milliammeter (mA).

Some other common symbols include the *incandescent lamp*, the *capacitor*, the *air-core coil*, the *iron-core coil*, the *chassis ground*, the *earth ground*, the *ac source*, the set of *terminals*, and the *black box* (general component or device), a rectangle with the designator written inside. These are shown in Fig. 4-4.

Schematic and Wiring Diagrams

Look back through the earlier chapters of this book and observe the electrical diagrams. These are all simple examples of how professionals would draw *schematic diagrams.* In a schematic diagram, the interconnection of the components is shown, but the actual values of the components are not necessarily indicated. You might see a diagram of a two-transistor audio amplifier, for example, with resistors and capacitors and coils and transistors, but without any data concerning the values or ratings of the components. This is a schematic diagram, but not a true *wiring diagram.* It gives the *scheme* for the circuit, but you can't *wire* the circuit and make it work, because there isn't enough information.



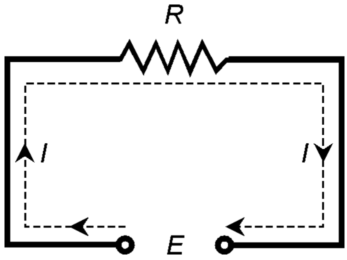
4-4 Schematic symbols for incandescent lamp (A), fixed capacitor (B), fixed inductor with air core (C), fixed inductor with laminated-iron core (D), chassis ground (E), earth ground (F), signal generator or source of alternating current (G), pair of terminals (H), and specialized component or device (I).

Suppose you want to build the circuit. You go to an electronics store to get the parts. What values of resistors should you buy? How about capacitors? What type of transistor will work best? Do you need to wind the coils yourself, or can you get ready-made coils? Are there test points or other special terminals that should be installed for the benefit of the technicians who might have to repair the amplifier? How many watts should the potentiometers be able to handle? All these things are indicated in a wiring diagram. You might have seen this kind of diagram in the back of the instruction manual for a hi-fi amplifier, a stereo tuner, or a television set. Wiring diagrams are especially useful when you want to build, modify, or repair an electronic device.

Voltage/Current/Resistance Circuits

Most dc circuits can be boiled down to three major components: a voltage source, a set of conductors, and a resistance. This is shown in Fig. 4-5. The voltage or EMF source is *E;* the current in the conductor is *I;* the resistance is *R.*

You already know that there is a relationship among these three quantities. If one of them changes, then one or both of the others will change. If you make the resistance smaller, the current will get larger. If you reduce the applied voltage, the current will also decrease. If the current in the circuit increases, the voltage across the resistor will increase. There is a simple arithmetic relationship among these three quantities.



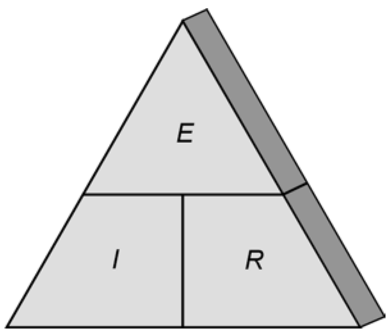
4-5 The basic elements of a dc circuit. The voltage is *E*, the current is *I*, and the resistance is *R*.

Ohm's Law

The interdependence among current, voltage, and resistance in dc circuits is called *Ohm's Law*, named after the scientist who supposedly first quantified it. Three formulas denote this law:

$$E = IR$$
$$I = E/R$$
$$R = E/I$$

You need only remember the first of these formulas in order to derive the others. The easiest way to remember it is to learn the abbreviations *E* for voltage, *I* for current, and *R* for resistance, and then remember that they appear in alphabetical order with the equal sign after the *E*. Sometimes the three symbols are arranged in the so-called Ohm's Law triangle, shown in Fig. 4-6. To find the value of a quantity, cover it up and read the positions of the others.



4-6 The Ohm's Law triangle. The voltage is *E*, the current is *I*, and the resistance is *R*. These quantities are expressed in volts, amperes, and ohms, respectively.

Remember that you must use units of volts, amperes, and ohms for the Ohm's Law formulas to yield a meaningful result! If you use, say, volts and microamperes to calculate a resistance, you cannot be sure of the units you'll end up with when you derive the final result. If the initial quantities are given in units other than volts, amperes, and ohms, convert to these units, and then calculate. After that, you can convert the calculated current, voltage, or resistance value to whatever size unit you want. For example, if you get 13,500,000 Ω as a calculated resistance, you might prefer to say that it's 13.5 MΩ.

Current Calculations

The first way to use Ohm's Law is to determine current in dc circuits. In order to find the current, you must know the voltage and the resistance, or be able to deduce them. Refer to the schematic diagram of Fig. 4-7. It consists of a dc voltage source, a voltmeter, some wire, an ammeter, and a calibrated, wide-range potentiometer.

Problem 4-1

Suppose that the dc generator in Fig. 4-7 produces 10 V and the potentiometer is set to a value of 10 Ω. What is the current?

This is solved by the formula $I = E/R$. Plug in the values for E and R; they are both 10, because the units are given in volts and ohms. Then $I = 10/10 = 1.0$ A.

Problem 4-2

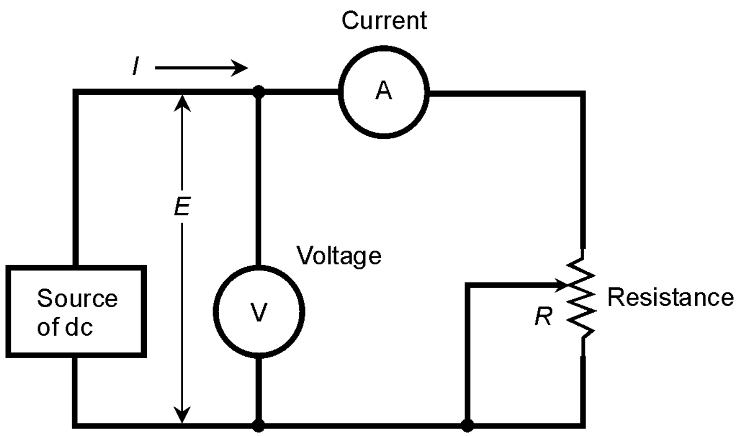
Imagine that dc generator in Fig. 4-7 produces 100 V and the potentiometer is set to 10 kΩ. What is the current?

First, convert the resistance to ohms: 10 kΩ = 10,000 Ω. Then plug the values in: $I = 100/10{,}000 = 0.01$ A. You might prefer to express this as 10 mA.

Problem 4-3

Suppose that dc generator in Fig. 4-7 is set to provide 88.5 V, and the potentiometer is set to 477 MΩ. What is the current?

This problem involves numbers that aren't exactly round, and one of them is huge. But you can use a calculator. First, change the resistance value to ohms, so you get 477,000,000 Ω. Then plug into the Ohm's Law formula: $I = E/R = 88.5 / 477{,}000{,}000 = 0.000000186$ A. It is more reasonable to express this as 0.186 μA or 186 nA.



4-7 A circuit for working Ohm's Law problems.

Voltage Calculations

The second application of Ohm's Law is to find unknown dc voltages when the current and the resistance are known. Let's work out some problems of this kind.

Problem 4-4

Suppose the potentiometer in Fig. 4-7 is set to 100 Ω, and the measured current is 10 mA. What is the dc voltage?

Use the formula $E = IR$. First, convert the current to amperes: 10 mA = 0.01 A. Then multiply: $E = 0.01 \times 100 = 1.0$ V. That's a little less than the voltage produced by a flashlight cell.

Problem 4-5

Adjust the potentiometer in Fig. 4-7 to a value of 157 kΩ, and suppose the current reading is 17.0 mA. What is the voltage of the source?

You must convert both the resistance and the current values to their proper units. A resistance of 157 kΩ is 157,000 Ω, and a current of 17.0 mA is 0.0170 A. Then $E = IR = 0.017 \times 157{,}000 =$ 2669 V = 2.669 kV. You should round this off to 2.67 kV. This is a dangerously high voltage.

Problem 4-6

Suppose you set the potentiometer in Fig. 4-7 so that the meter reads 1.445 A, and you observe that the potentiometer scale shows 99 Ω. What is the voltage?

These units are both in their proper form. Therefore, you can plug them right in and use your calculator: $E = IR = 1.445 \times 99 = 143.055$ V. This can and should be rounded off—but to what extent? This is a good time to state an important rule that should be followed in all technical calculations.

The Rule of Significant Figures

Competent engineers and scientists go by the *rule of significant figures*, also called the *rule of significant digits*. After completing a calculation, you should always round the answer off to the *least* number of digits given in the input data numbers.

If you follow this rule in Problem 4-6, you must round off the answer to two significant digits, getting 140 V, because the resistance (99 Ω) is only specified to that level of accuracy. If the resistance were given as 99.0 Ω, then you would round off the answer to 143 V. If the resistance were given as 99.00 Ω, then you could state the answer as 143.1 V. However, any further precision in the resistance value would not entitle you to go to any more digits in your answer, unless the current were specified to more than four significant figures.

This rule takes some getting used to if you haven't known about it or practiced it before. But after a while, it will become a habit.

Resistance Calculations

Ohms' Law can be used to find a resistance between two points in a dc circuit when the voltage and the current are known.

Problem 4-7

If the voltmeter in Fig. 4-7 reads 24 V and the ammeter shows 3.0 A, what is the resistance of the potentiometer?

Use the formula $R = E/I$, and plug in the values directly, because they are expressed in volts and amperes: $R = 24/3.0 = 8.0\ \Omega$. Note that you can specify this value to two significant figures, the 8 and the 0, rather than saying simply 8 Ω. This is because you are given both the voltage and the current to two significant figures. If the ammeter reading had been given as 3 A, you would only be entitled to express the answer as 8 Ω, to one significant digit. The digit 0 can be, and often is, just as important in calculations as any of the other digits 1 through 9.

Problem 4-8

What is the value of the resistance in Fig. 4-7 if the current is 18 mA and the voltage is 229 mV?

First, convert these values to amperes and volts. This gives $I = 0.018$ A and $E = 0.229$ V. Then plug into the equation: $R = E/I = 0.229/0.018 = 13\ \Omega$.

Problem 4-9

Suppose the ammeter in Fig. 4-7 reads 52 μA and the voltmeter indicates 2.33 kV. What is the resistance?

Convert to amperes and volts, getting $I = 0.000052$ A and $E = 2330$ V. Then plug into the formula: $R = E/I = 2330/0.000052 = 45{,}000{,}000\ \Omega = 45\ \text{M}\Omega$.

Power Calculations

You can calculate the power P, in watts, in a dc circuit such as that shown in Fig. 4-7, by using the formula $P = EI$. This formula tells us that the power in watts is the product of the voltage in volts and the current in amperes. If you are not given the voltage directly, you can calculate it if you know the current and the resistance.

Recall the Ohm's Law formula for obtaining voltage: $E = IR$. If you know I and R but you don't know E, you can get the power P this way:

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

Suppose you're given only the voltage and the resistance. Remember the Ohm's Law formula for obtaining current: $I = E/R$. Therefore:

$$P = EI = E(E/R) = E^2/R$$

Problem 4-10

Suppose that the voltmeter in Fig. 4-7 reads 12 V and the ammeter shows 50 mA. What is the power dissipated by the potentiometer?

Use the formula $P = EI$. First, convert the current to amperes, getting $I = 0.050$ A. (Note that the last 0 counts as a significant digit.) Then multiply by 12 V, getting $P = EI = 12 \times 0.050 = 0.60$ W.

Problem 4-11

If the resistance in the circuit of Fig. 4-7 is 999 Ω and the voltage source delivers 3 V, what is the power dissipated by the potentiometer?

Use the formula $P = E^2/R = 3 \times 3/999 = 9/999 = 0.009$ W = 9 mW. You are justified in going to only one significant figure here.

Problem 4-12

Suppose the resistance in Fig. 4-7 is 47 kΩ and the current is 680 mA. What is the power dissipated by the potentiometer?

Use the formula $P = I^2R$, after converting to ohms and amperes. Then $P = 0.680 \times 0.680 \times 47{,}000 = 22{,}000$ W = 22 kW. (This is an unrealistic state of affairs: an ordinary potentiometer, such as the type you would use as the volume control in a radio, dissipating 22 kW, several times more than a typical household!)

Problem 4-13

How much voltage would be necessary to drive 680 mA through a resistance of 47 kΩ, as is described in the previous problem?

Use Ohm's Law to find the voltage: $E = IR = 0.680 \times 47{,}000 = 32{,}000$ V = 32 kV. That's the level of voltage you'd expect to find on a major utility power line, or in a high-power tube-type radio broadcast transmitter.

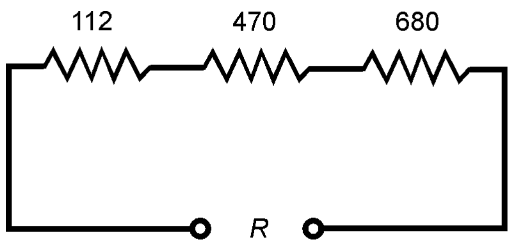
Resistances in Series

When you place resistances in series, their ohmic values add together to get the total resistance. This is easy to imagine, and it's easy to remember!

Problem 4-14

Suppose resistors with the following values are connected in series, as shown in Fig. 4-8: 112 Ω, 470 Ω, and 680 Ω. What is the total resistance of the series combination?

Simply add up the values, getting a total of 112 + 470 + 680 = 1262 Ω. You might round this off to 1260 Ω. It depends on the *tolerances* of the resistors—how precise their actual values are to the ones specified by the manufacturer.



4-8 Three resistors in series. Illustration for Problem 4-14. Resistance values are in ohms.

Resistances in Parallel

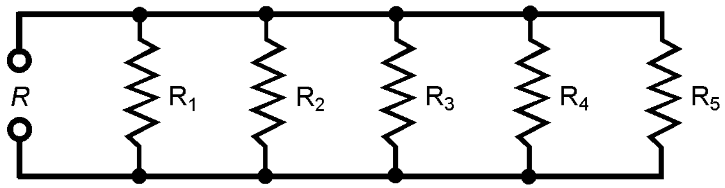
When resistances are placed in parallel, they behave differently than they do in series. One way to look at resistances in parallel is to consider them as conductances instead. In parallel, conductances add up directly, just as resistances add up in series. If you change all the ohmic values to siemens, you can add these figures up and convert the final answer back to ohms.

The symbol for conductance is *G*. This figure, in siemens, is related to the resistance *R*, in ohms, by these formulas, which you learned in Chap. 2:

$$G = 1/R$$
$$R = 1/G$$

Problem 4-15

Consider five resistors in parallel. Call them R_1 through R_5, and call the total resistance *R* as shown in Fig. 4-9. Let the resistance values be as follows: $R_1 = 100\ \Omega$, $R_2 = 200\ \Omega$, $R_3 = 300\ \Omega$, $R_4 = 400\ \Omega$, and $R_5 = 500\ \Omega$. What is the total resistance, *R*, of this parallel combination?



4-9 Five resistors of values R_1 through R_5, connected in parallel, produce a net resistance *R*. Illustration for Problems 4-15 and 4-16.

Converting the resistances to conductance values, you get: $G_1 = 1/100 = 0.01$ S, $G_2 = 1/200 = 0.005$ S, $G_3 = 1/300 = 0.00333$ S, $G_4 = 1/400 = 0.0025$ S, and $G_5 = 1/500 = 0.002$ S. Adding these gives $G = 0.01 + 0.005 + 0.00333 + 0.0025 + 0.002 = 0.0228$ S. The total resistance is therefore $R = 1/G = 1/0.0228 = 43.8\ \Omega$.

Problem 4-16

Suppose you have five resistors, called R_1 through R_5, connected in parallel as shown in Fig. 4-9. Suppose all the resistances, R_1 through R_5, are 4.70 kΩ. What is the total resistance, *R*, of this combination?

When you have two or more resistors connected in parallel and their resistances are all the same, the total resistance is equal to the resistance of any one component divided by the number of components. In this example, convert the resistance of any single resistor to 4700 Ω, and then divide this by 5. Thus, you can see that the total resistance is 4700/5 = 940 Ω.

In a situation like this, where you have a bunch of resistors connected together to operate as a single unit, the total resistance is sometimes called the *net resistance*. Take note, too, that R is not italicized when it means *resistor*, but *R* is italicized when it means *resistance!*

Division of Power

When combinations of resistances are connected to a source of voltage, they draw current. You can figure out how much current they draw by calculating the total resistance of the combination, and then considering the network as a single resistor.

If the resistors in the network all have the same ohmic value, the power from the source is evenly distributed among them, whether they are hooked up in series or in parallel. For example, if there are eight identical resistors in series with a battery, the network consumes a certain amount of power, each resistor bearing $1/8$ of the load. If you rearrange the circuit so that the resistors are in parallel, the circuit will dissipate a certain amount of power (a lot more than when the resistors were in series), but again, each resistor will handle $1/8$ of the total power load.

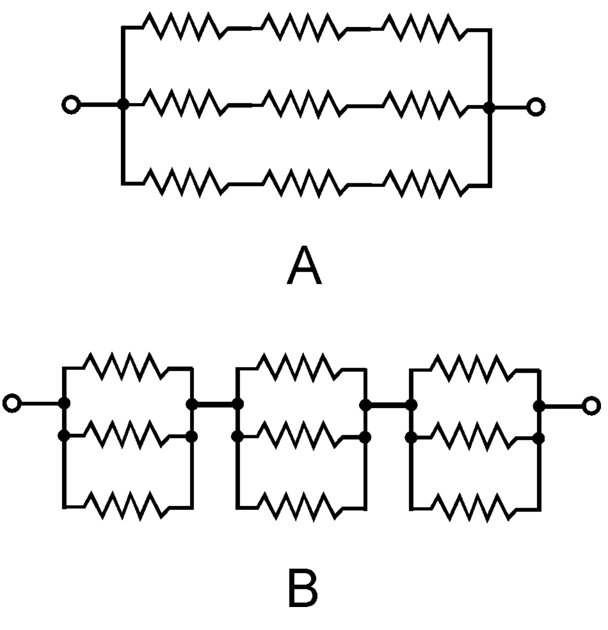
If the resistances in the network do not all have identical ohmic values, they divide up the power unevenly. Situations like this are discussed in the next chapter.

Resistances in Series-Parallel

Sets of resistors, all having identical ohmic values, can be connected together in parallel sets of series networks, or in series sets of parallel networks. By doing this, the total power-handling capacity of the resistance can be greatly increased over that of a single resistor.

Sometimes, the total resistance of a *series-parallel network* is the same as the value of any one of the resistors. This is always true if the components are identical, and are in a network called an *n-by-n matrix*. That means, when n is a whole number, there are n parallel sets of n resistors in series (Fig. 4-10A), or else there are n series sets of n resistors in parallel (Fig. 4-10B). Either arrangement gives the same practical result.

Engineers and technicians sometimes use series-parallel networks to obtain resistances with large power-handling capacity. A series-parallel array of n by n resistors will have n^2 times that of a single resistor. Thus, a 3×3 series-parallel matrix of 2 W resistors can handle up to $3^2 \times 2 = 9 \times$



4-10 Series-parallel resistances. At A, sets of series resistors are connected in parallel. At B, sets of parallel resistances are connected in series. These examples show symmetrical n-by-n matrices with $n = 3$.

2 = 18 W, for example. A 10 × 10 array of 1-W resistors can dissipate up to 100 W. The total power-handling capacity is multiplied by the total number of resistors in the matrix. But this is true only if all the resistors have the same ohmic values, and the same power-dissipation ratings.

It is unwise to build series-parallel arrays from resistors with different ohmic values or power ratings. If the resistors have values and/or ratings that are even a little nonuniform, one of them might be subjected to more current than it can withstand, and it will burn out. Then the current distribution in the network can change so a second component fails, and then a third. It's hard to predict the current and power distribution in an array when its resistor values are all different.

If you need a resistance with a certain power-handling capacity, you must be sure the network can handle at least that much power. If a 50-W rating is required, and a certain combination will handle 75 W, that's fine. But it isn't good enough to build a circuit that will handle only 48 W. Some extra tolerance, say 10 percent over the minimum rating needed, is good, but it's silly to make a 500-W network using far more resistors than necessary, unless that's the only convenient combination given the parts available.

Nonsymmetrical series-parallel networks, made up from identical resistors, can increase the power-handling capability over that of a single resistor. But in these cases, the total resistance is not the same as the value of the single resistors. The overall power-handling capacity is always multiplied by the total number of resistors, whether the network is symmetrical or not, provided all the ohmic values are identical. In engineering work, cases sometimes arise where nonsymmetrical networks fit the need.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct answers. The answers are in the back of the book.

1. Suppose you double the voltage in a simple dc circuit, and cut the resistance in half. The current will
   - (a) become four times as great.
   - (b) become twice as great.
   - (c) stay the same as it was before.
   - (d) become half as great.

2. You can expect to find a wiring diagram
   - (a) on a sticker on the back of a television receiver.
   - (b) in an advertisement for an electric oven.
   - (c) in the service/repair manual for a two-way radio.
   - (d) in the photograph of the front panel of a stereo hi-fi tuner.

For questions 3 through 11, please refer to Fig. 4-7. Remember to take significant figures into account when completing your calculations!

3. Given a dc voltage source delivering 24 V and a resistance of 3.3 kΩ, what is the current?
   - (a) 0.73 A
   - (b) 138 A
   - (c) 138 mA
   - (d) 7.3 mA

4. Suppose the resistance is 472 Ω, and the current is 875 mA. The source voltage must therefore be
   (a) 413 V.
   (b) 0.539 V.
   (c) 1.85 V.
   (d) none of the above.

5. Suppose the dc voltage is 550 mV and the current is 7.2 mA. Then the resistance is
   (a) 0.76 Ω.
   (b) 76 Ω.
   (c) 0.0040 Ω.
   (d) none of the above.

6. Given a dc voltage source of 3.5 kV and a resistance of 220 Ω, what is the current?
   (a) 16 mA
   (b) 6.3 mA
   (c) 6.3 A
   (d) None of the above

7. Suppose the resistance is 473,332 Ω, and the current flowing through it is 4.4 mA. The best expression for the voltage of the source is
   (a) 2082 V.
   (b) 110 kV.
   (c) 2.1 kV.
   (d) 2.08266 kV.

8. A source delivers 12 V and the current is 777 mA. The best expression for the resistance is
   (a) 15 Ω.
   (b) 15.4 Ω.
   (c) 9.3 Ω.
   (d) 9.32 Ω.

9. Suppose the voltage is 250 V and the current is 8.0 mA. The power dissipated by the potentiometer is
   (a) 31 mW.
   (b) 31 W.
   (c) 2.0 W.
   (d) 2.0 mW.

10. Suppose the voltage from the source is 12 V and the potentiometer is set for 470 Ω. The power dissipated in the resistance is approximately
   (a) 310 mW.
   (b) 25.5 mW.
   (c) 39.2 W.
   (d) 3.26 W.

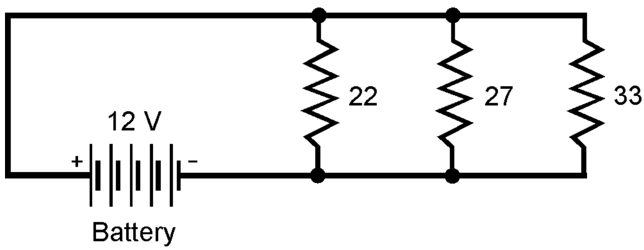
11. If the current through the potentiometer is 17 mA and its resistance is set to 1.22 kΩ, what is the power dissipated by it?
    (a) 0.24 μW
    (b) 20.7 W
    (c) 20.7 mW
    (d) 350 mW

12. Suppose six resistors are hooked up in series, and each of them has a value of 540 Ω. What is the resistance across the entire combination?
    (a) 90 Ω
    (b) 3.24 kΩ
    (c) 540 Ω
    (c) None of the above

13. If four resistors are connected in series, each with a value of 4.0 kΩ, the total resistance is
    (a) 1 kΩ.
    (b) 4 kΩ.
    (c) 8 kΩ.
    (d) 16 kΩ.

14. Suppose you have three resistors in parallel, each with a value of 0.069 MΩ. Then the total resistance is
    (a) 23 Ω.
    (b) 23 kΩ.
    (c) 204 Ω.
    (d) 0.2 MΩ.

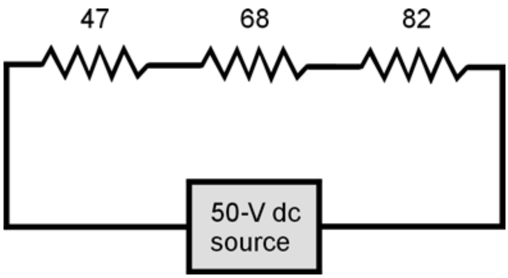
15. Imagine three resistors in parallel, with values of 22 Ω, 27 Ω, and 33 Ω. If a 12-V battery is connected across this combination, as shown in Fig. 4-11, what is the current drawn from the battery?
    (a) 1.4 A
    (b) 15 mA
    (c) 150 mA
    (d) 1.5 A



4-11 Illustration for Quiz Question 15. Resistance values are in ohms.

16. Imagine three resistors, with values of 47 Ω, 68 Ω, and 82 Ω, connected in series with a 50-V dc generator, as shown in Fig. 4-12. The total power consumed by this network of resistors is

(a) 250 mW.
(b) 13 mW.
(c) 13 W.
(d) impossible to determine from the data given.



4-12 Illustration for Quiz Question 16. Resistance values are in ohms.

17. Suppose you have an unlimited supply of 1-W, 100-Ω resistors. You need to get a 100-Ω, 10-W resistor. This can be done most cheaply by means of a series-parallel matrix of

(a) 3×3 resistors.
(b) 4×3 resistors.
(c) 4×4 resistors.
(d) 2×5 resistors.

18. Suppose you have an unlimited supply of 1-W, 1000-Ω resistors, and you need a 500-Ω resistance rated at 7 W or more. This can be done by assembling

(a) four sets of two resistors in series, and connecting these four sets in parallel.
(b) four sets of two resistors in parallel, and connecting these four sets in series.
(c) a 3×3 series-parallel matrix of resistors.
(d) a series-parallel matrix, but something different than those described above.

19. Suppose you have an unlimited supply of 1-W, 1000-Ω resistors, and you need to get a 3000-Ω, 5-W resistance. The best way is to

(a) make a 2×2 series-parallel matrix.
(b) connect three of the resistors in parallel.
(c) make a 3×3 series-parallel matrix.
(d) do something other than any of the above.

20. Good engineering practice usually requires that a series-parallel resistive network be assembled

(a) from resistors that are all different.
(b) from resistors that are all identical.
(c) from a series combination of resistors in parallel but not from a parallel combination of resistors in series.
(d) from a parallel combination of resistors in series, but not from a series combination of resistors in parallel.

5

CHAPTER

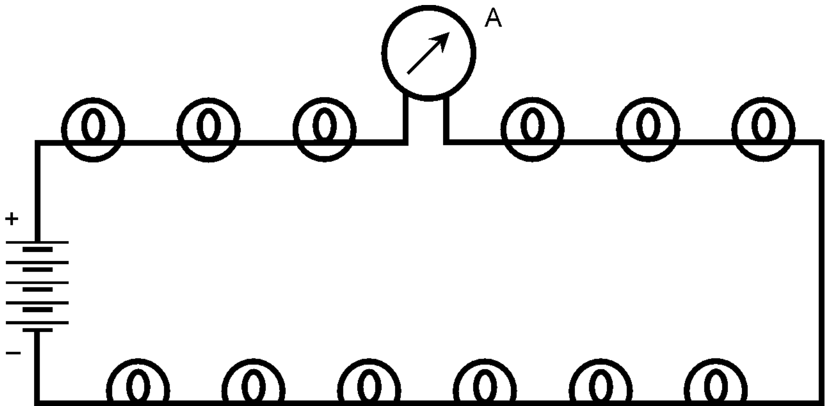
Direct-Current Circuit Analysis

IN THIS CHAPTER, YOU'LL LEARN MORE ABOUT DC CIRCUITS AND HOW THEY BEHAVE UNDER VARIOUS conditions. These principles apply to most ac utility circuits as well.

Current through Series Resistances

Have you ever used those tiny holiday lights that come in strings? If one bulb burns out, the whole set of bulbs goes dark. Then you have to find out which bulb is bad, and replace it to get the lights working again. Each bulb works with something like 10 V; there are about a dozen bulbs in the string. You plug in the whole bunch and the 120-V utility mains drive just the right amount of current through each bulb.

In a series circuit, such as a string of light bulbs (Fig. 5-1), the current at any given point is the same as the current at any other point. The ammeter, A, is shown in the line between two of the bulbs. If it were moved anywhere else along the current path, it would indicate the same current.



5-1 Light bulbs in series, with an ammeter (A) in the circuit.

This is true in any series dc circuit, no matter what the components actually are, and regardless of whether or not they all have the same resistance.

If the bulbs in Fig. 5-1 had different resistances, some of them would consume more power than others. In case one of the bulbs in Fig. 5-1 burns out, and its socket is then shorted out instead of filled with a replacement bulb, the current through the whole chain will increase, because the overall resistance of the string will go down. This will force each of the remaining bulbs to carry more current, and pretty soon another bulb would burn out because of the excessive current. If it, too, were replaced with a short circuit, the current would be increased still further. A third bulb would blow out almost right away thereafter.

Voltages across Series Resistances

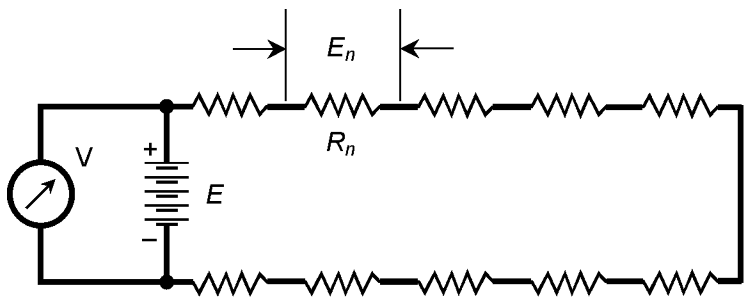
The bulbs in the string of Fig. 5-1, being all the same, each get the same amount of voltage from the source. If there are a dozen bulbs in a 120-V circuit, each bulb has a potential difference of 10 V across it. This will remain true even if the bulbs are replaced with brighter or dimmer ones, as long as all the bulbs in the string are identical.

Look at the schematic diagram of Fig. 5-2. Each resistor carries the same current. Each resistance R_n has a potential difference E_n across it equal to the product of the current and the resistance of that particular resistor. The voltages E_n are in series, like cells in a battery, so they add together. What if the voltages across all the resistors added up to something more or less than the supply voltage, E? Then there would have to be a "phantom EMF" someplace, adding or taking away voltage. But that's impossible. Voltage cannot come out of nowhere!

Look at this another way. The voltmeter V in Fig. 5-2 shows the voltage E of the battery, because the meter is hooked up across the battery. The voltmeter V also shows the sum of the voltages E_n across the set of resistances, because it's connected across the whole combination. The meter says the same thing whether you think of it as measuring the battery voltage E or as measuring the sum of the voltages E_n across the series combination of resistances. Therefore, E is equal to the sum of the voltages E_n.

How do you find the voltage across any particular resistance R_n in a circuit like the one in Fig. 5-2? Remember Ohm's Law for finding voltage: $E = IR$. Remember, too, that you must use volts, ohms, and amperes when making calculations.

In order to find the current in the circuit, I, you need to know the total resistance and the supply voltage; then $I = E/R$. First find the current in the whole circuit; then find the voltage across any particular resistor.



5-2 Analysis of voltages in a series circuit.

Problem 5-1

In Fig. 5-2, there are 10 resistors. Five of them have values of 10 Ω, and the other five have values of 20 Ω. The power source is 15-V dc. What is the voltage across any one of the 10-Ω resistors? Across any one of the 20-Ω resistors?

First, find the total resistance: $R = (10 \times 5) + (20 \times 5) = 50 + 100 = 150$ Ω. Then find the current: $I = E/R = 15/150 = 0.10$ A. This is the current through each of the resistances in the circuit.

- If $R_n = 10$ Ω, then $E_n = IR_n = 0.1 \times 10 = 1.0$ V.
- If $R_n = 20$ Ω, then $E_n = IR_n = 0.1 \times 20 = 2.0$ V.

Let's check to be sure all of these voltages add up to the supply voltage. There are five resistors with 1.0 V across each, for a total of 5.0 V; there are also five resistors with 2.0 V across each, for a total of 10 V. So the sum of the voltages across the resistors is 5.0 + 10 = 15 V.

Problem 5-2

In the circuit of Fig. 5-2, what will happen to the voltages across the resistances if one of the 20-Ω resistances is replaced with a short circuit?

In this case the total resistance becomes $R = (10 \times 5) + (20 \times 4) = 50 + 80 = 130$ Ω. The current is therefore $I = E/R = 15/130 = 0.12$ A. This is the current at any point in the circuit, rounded off to two significant figures.

The voltage E_n across any of the 10-Ω resistances R_n is equal to IR_n, which is $0.12 \times 10 = 1.2$ V. The voltage E_n across any of the 20-Ω resistances R_n is equal to IR_n, which is $0.12 \times 20 = 2.4$ V. Checking the total voltage, add $(5 \times 1.2) + (4 \times 2.4) = 6.0 + 9.6 = 15.6$ V. This rounds off to 16 V when we cut it down to two significant figures.

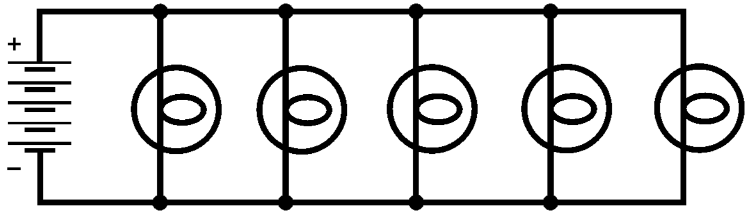
A "Rounding-Off Bug"

Compare the result for total voltage in Problem 5-2 with the result for total voltage in Problem 5-1. What is going on here? Where does the extra volt come from in the second calculation? Certainly, shorting out one of the resistances cannot cause the battery voltage to change!

This is an example of what can happen when you round off to a certain number of significant figures after calculating the value of some parameter X in a circuit, then change a different parameter Y in the circuit, and finally calculate the value of X again, rounding off to the same number of significant digits as you did the first time. The discrepancy is the result of a "rounding-off bug."

If this bug bothers you (and it should), keep all the digits your calculator will hold while you go through the solution process for Problem 5-2. The current in the circuit, as obtained by means of a calculator that can show 10 digits, should come out as 0.115384615 A. When you find the voltages across all the resistances R_n, accurate to all these extra digits, and then add them up, you'll get a final rounded-off voltage of 15 V.

This example shows why it is a good idea to wait until you get the final answer in a calculation, or set of calculations, involving a particular circuit before you round off to the allowed number of significant digits. Rounding-off bugs of the sort we have just seen can be more than mere annoyances. They are easy to overlook, but they can generate large errors in *iterative processes* involving calculations that are done over and over.



5-3 Light bulbs in parallel.

Voltage across Parallel Resistances

Imagine a set of ornamental light bulbs connected in parallel (Fig. 5-3). This is the method used for outdoor holiday lighting or for bright indoor lighting. It's easier to repair a parallel-wired string of such lights if one bulb should burn out than it is to fix a series-wired string. And in the parallel configuration, the failure of one bulb does not cause total system failure.

In a parallel circuit, the voltage across each component is equal to the supply or battery voltage. The current drawn by each component depends only on the resistance of that particular device. In this sense, the components in a parallel-wired circuit operate independently, as opposed to the series-wired circuit in which they all interact.

If any one branch of a parallel circuit opens up, is disconnected, or is removed, the conditions in the other branches do not change. If new branches are added, assuming the power supply can handle the load, conditions in previously existing branches are not affected.

Currents through Parallel Resistances

Refer to the schematic diagram of Fig. 5-4. The resistances are called R_n. The total parallel resistance in the circuit is R. The battery voltage is E. The current in any particular branch n, containing resistance R_n, is measured by ammeter A and is called I_n. The sum of all the currents I_n is equal to the total current, I, drawn from the battery. The current is divided up in the parallel circuit in a manner similar to the way that voltage is divided up in a series circuit.

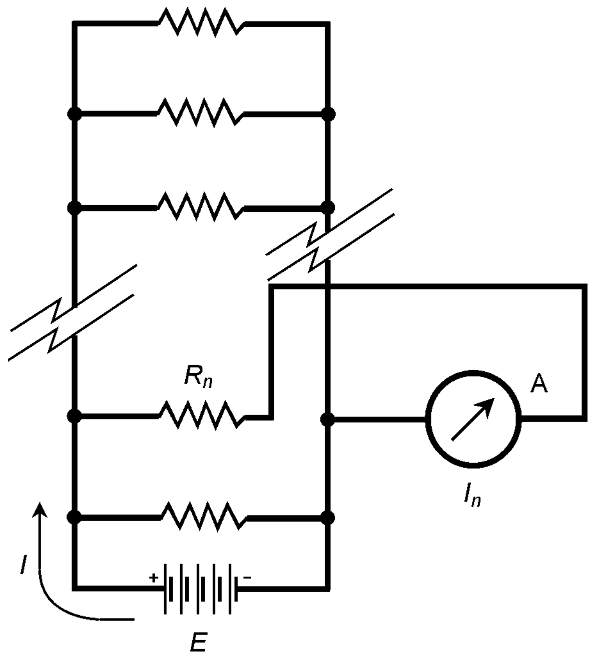
Conventional Current

Have you noticed that the direction of current flow in Fig. 5-4 is portrayed as outward from the positive battery terminal? Don't electrons, which are the actual charge carriers in a wire, flow out of the minus terminal of a battery? Yes, that's true; but scientists consider *theoretical current*, more often called *conventional current* (because it is defined by convention), to flow from positive to negative voltage points, rather than from negative to positive.

Problem 5-3

Suppose that the battery in Fig. 5-4 delivers 12 V. Further suppose that there are 12 resistors, each with a value of 120 Ω in the parallel circuit. What is the total current, I, drawn from the battery?

First, find the total resistance. This is easy, because all the resistors have the same value. Just divide $R_n = 120$ by 12 to get $R = 10\ \Omega$. Then the current can be found by Ohm's Law: $I = E/R = 12/10 = 1.2$ A.



5-4 Analysis of current in a parallel circuit.

Problem 5-4

In the circuit of Fig. 5-4, what does the ammeter say?

This involves finding the current in any given branch. The voltage is 12 V across every branch, and $R_n = 120\ \Omega$. Therefore I_n, the ammeter reading, is found by Ohm's Law: $I_n = E/R_n = 12/120 = 0.10$ A.

Because this is a parallel circuit, all of the branch currents I_n should add up to get the total current, I. There are 12 identical branches, each carrying 0.10 A; therefore the total current is $0.10 \times 12 = 1.2$ A. It checks out.

Problem 5-5

Suppose three resistors are in parallel across a battery that supplies $E = 12$ V. The resistances are $R_1 = 22\ \Omega$, $R_2 = 47\ \Omega$, and $R_3 = 68\ \Omega$. These resistances carry currents I_1, I_2, and I_3, respectively. What is the current, I_3, through R_3?

This problem is solved by means of Ohm's Law as if R_3 is the only resistance in the circuit. There's no need to worry about the parallel combination. The other branches do not affect I_3. Thus $I_3 = E/R_3 = 12/68 = 0.18$ A.

Problem 5-6

What is the total current drawn by the circuit described in Problem 5-5?

There are two ways to go at this. One method involves finding the total resistance, R, of R_1, R_2, and R_3 in parallel, and then calculating I based on R. Another way is to find the currents through R_1, R_2, and R_3 individually, and then add them up.

Using the first method, first change the resistances R_n into conductances G_n. This gives $G_1 = 1/R_1 = 1/22 = 0.04545$ S, $G_2 = 1/R_2 = 1/47 = 0.02128$ S, and $G_3 = 1/R_3 = 1/68 = 0.01471$ S. Adding these gives $G = 0.08144$ S. The resistance is therefore $R = 1/G = 1/0.08144 = 12.279\ \Omega$. Use Ohm's Law to find $I = E/R = 12/12.279 = 0.98$ A. Note that extra digits are used throughout the calculation, rounding off only at the end.

Now let's try the other method. Find $I_1 = E/R_1 = 12/22 = 0.5455$ A, $I_2 = E/R_2 = 12/47 = 0.2553$ A, and $I_3 = E/R_3 = 12/68 = 0.1765$ A. Adding these gives $I = I_1 + I_2 + I_3 = 0.5455 + 0.2553 + 0.1765 = 0.9773$ A, which rounds off to 0.98 A.

Power Distribution in Series Circuits

When calculating the power in a circuit containing resistors in series, all you need to do is find out the current, I, that the circuit is carrying. Then it's easy to calculate the power P_n dissipated by any one of the resistances R_n, based on the formula $P_n = I^2R_n$.

Problem 5-7

Suppose we have a series circuit with a supply of 150 V and three resistances: $R_1 = 330\ \Omega$, $R_2 = 680\ \Omega$, and $R_3 = 910\ \Omega$. What is the power dissipated by R_2?

First, find the current that flows through the circuit. Calculate the total resistance first. Because the resistors are in series, the total is $R = 330 + 680 + 910 = 1920\ \Omega$. The current is $I = 150/1920 = 0.07813$ A. The power dissipated by R_2 is therefore $P_2 = I^2R_2 = 0.07813 \times 0.07813 \times 680 = 4.151$ W. Round this off to three significant digits, because that's all we have in the data, to obtain 4.15 W.

The total wattage dissipated in a series circuit is equal to the sum of the wattages dissipated in each resistance.

Problem 5-8

Calculate the total dissipated power P in the circuit of Problem 5-7 by two different methods.

First, let's figure out the power dissipated by each of the three resistances separately, and then add the figures up. The power P_2 is already known. Let's use all the significant digits we have while we calculate. Thus, as found in Problem 5-7, $P_2 = 4.151$ W. Recall that the current is $I = 0.07813$ A. Then $P_1 = 0.07813 \times 0.07813 \times 330 = 2.014$ W, and $P_3 = 0.07813 \times 0.07813 \times 910 = 5.555$ W. Adding the three power figures gives us $P = P_1 + P_2 + P_3 = 2.014 + 4.151 + 5.555 = 11.720$ W. We should round this off to 11.7 W.

The second method is to find the total series resistance and then calculate the power. The series resistance is $R = 1920\ \Omega$, as found in Problem 5-7. Then $P = I^2R = 0.07813 \times 0.07813 \times 1920 = 11.72$ W. Again, we should round this to 11.7 W.

Power Distribution in Parallel Circuits

When resistances are wired in parallel, they each consume power according to the same formula, $P = I^2R$. But the current is not the same in each resistance. An easier method to find the power P_n dissipated by each of the various resistances R_n is to use the formula $P_n = E^2/R_n$, where E is the voltage of the supply or battery. This voltage is the same across every branch resistance in a parallel circuit.

Problem 5-9

Suppose a dc circuit contains three resistances $R_1 = 22\ \Omega$, $R_2 = 47\ \Omega$, and $R_3 = 68\ \Omega$ across a battery that supplies a voltage of $E = 3.0$ V. Find the power dissipated by each resistance.

Let's find the square of the supply voltage, E^2, first. We'll be needing this figure often: $E^2 = 3.0 \times 3.0 = 9.0$. Then the wattages dissipated by resistances R_1, R_2, and R_3 respectively are $P_1 = 9.0/22 = 0.4091$ W, $P_2 = 9.0/47 = 0.1915$ W, and $P_3 = 9.0/68 = 0.1324$ W. These should be rounded off to $P_1 = 0.41$ W, $P_2 = 0.19$ W, and $P_3 = 0.13$ W. (But let's remember the values to four significant figures for the next problem!)

In a parallel circuit, the total dissipated wattage is equal to the sum of the wattages dissipated by the individual resistances.

Problem 5-10

Find the total consumed power of the resistor circuit in Problem 5-9 using two different methods.

The first method involves adding P_1, P_2, and P_3. Let's use the four-significant-digit values to avoid the possibility of encountering the rounding-off bug. The total power thus calculated is $P = 0.4091 + 0.1915 + 0.1324 = 0.7330$ W. Now that we've finished the calculation, we should round it off to 0.73 W.

The second method involves finding the net resistance R of the parallel combination. You can do this calculation yourself. Determining it to four significant digits, you should get a net resistance of $R = 12.28\ \Omega$. Then $P = E^2/R = 9.0/12.28 = 0.7329$ W. Now that the calculation is done, this can be rounded to 0.73 W.

It's the Law!

In electricity and electronics, dc circuit analysis can be made easier if you are acquainted with certain axioms, or *laws.* Here they are:

- The current in a series circuit is the same at every point along the way.
- The voltage across any resistance in a parallel combination of resistances is the same as the voltage across any other resistance, or across the whole set of resistances.
- The voltages across resistances in a series circuit always add up to the supply voltage.
- The currents through resistances in a parallel circuit always add up to the total current drawn from the supply.
- The total wattage consumed in a series or parallel circuit is always equal to the sum of the wattages dissipated in each of the resistances.

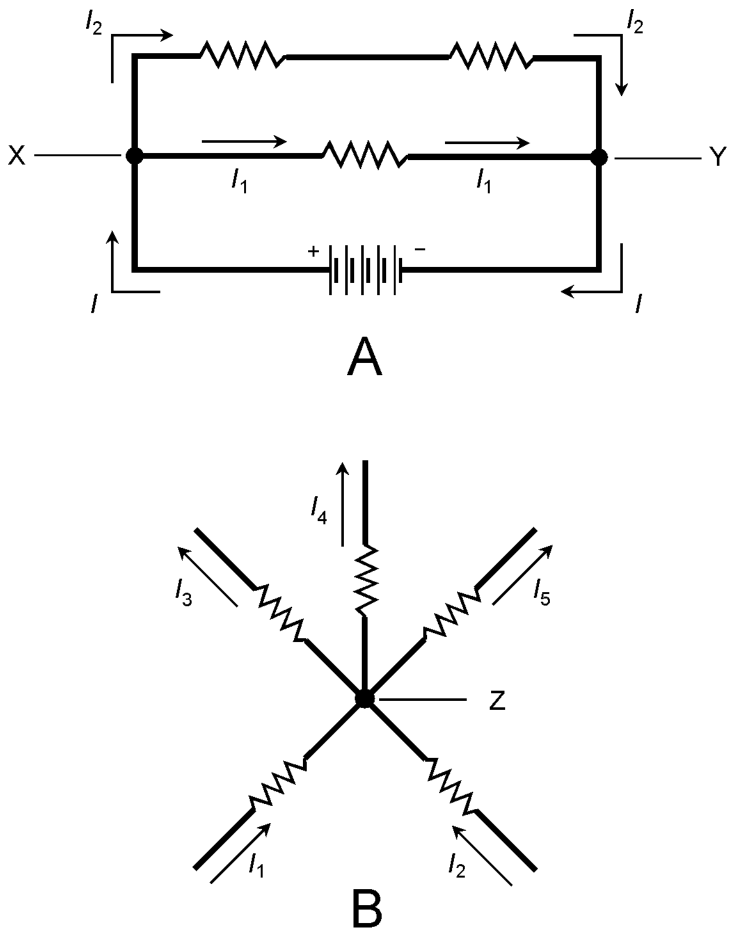
Now, let's get acquainted with two of the most famous laws that govern dc circuits. These rules are broad and sweeping, and they make it possible to analyze complicated series-parallel dc networks.

Kirchhoff's First Law

The physicist Gustav Robert Kirchhoff (1824–1887) was a researcher and experimentalist in a time when little was understood about how electric currents flow. Nevertheless, he used certain common-sense notions to deduce two important properties of dc circuits.

Kirchhoff reasoned that dc ought to behave something like water in a network of pipes, and that the current going into any point ought to be the same as the current going out of that point. This, Kirchhoff thought, must be true for any point in a circuit, no matter how many branches lead into or out of the point.

Two examples of this principles are shown in Fig. 5-5. Examine illustration A. At point X, I, the current going in, equals $I_1 + I_2$, the current going out. At point Y, $I_2 + I_1$, the current going in, equals I, the current going out. Now look at illustration B. In this case, at point Z, the current $I_1 + I_2$ going in is equal to the current $I_3 + I_4 + I_5$ going out. These are examples of *Kirchhoff's First Law.* We can also call it *Kirchhoff's Current Law* or the principle of *conservation of current.*



5-5 Kirchhoff's First Law. At A, the current into point X or point Y is the same as the current out of that point. That is, $I = I_1 + I_2$. At B, the current into point Z equals the current flowing out of point Z. That is, $I_1 + I_2 = I_3 + I_4 + I_5$. Illustration for Quiz Questions 13 and 14.

Problem 5-11

Refer to Fig. 5-5A. Suppose all three resistors have values of 100 Ω, and that $I_1 = 2.0$ A and $I_2 = 1.0$ A. What is the battery voltage?

First, find the current I drawn from the battery: $I = I_1 + I_2 = 2.0 + 1.0 = 3.0$ A. Next, find the resistance of the entire network. The two 100-Ω resistances in series give a value of 200 Ω, and this is in parallel with 100 Ω. You can do the calculations and find that the total resistance, R, connected across the battery is 66.67 Ω. Then $E = IR = 66.67 \times 3.0 = 200$ V.

Problem 5-12

In Fig. 5-5B, suppose each of the two resistors below point Z has a value of 100 Ω, and all three resistors above point Z have values of 10.0 Ω. Suppose the current through each 100-Ω resistor is 500 mA. What is the current through any one of the 10.0-Ω resistors, assuming that the current through all three 10.0-Ω resistors is the same? What is the voltage across any one of the three 10.0-Ω resistors?

The total current into point Z is 500 mA + 500 mA = 1.00 A. This is divided equally among the three 10-Ω resistors. Therefore, the current through any one of them is 1.00/3 A = 0.333 A. The voltage across any one of the 10.0-Ω resistors can thus found by Ohm's Law: $E = IR = 0.333 \times 10.0 = 3.33$ V.

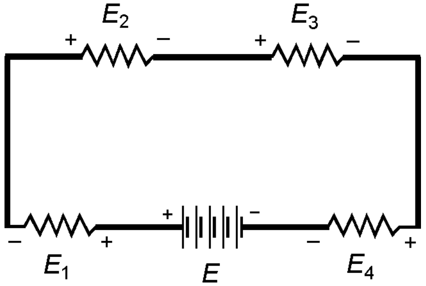
Kirchhoff's Second Law

The sum of all the voltages, as you go around a circuit from some fixed point and return there from the opposite direction, and taking polarity into account, is always zero. Does this seem counterintuitive? Let's think about it a little more carefully.

What Kirchhoff was expressing, when he wrote his second law, is the principle that voltage cannot appear out of nowhere, nor can it vanish. All the potential differences must ultimately cancel each other out in any closed dc circuit, no matter how complicated that circuit happens to be. This is *Kirchhoff's Second Law.* We can also call it *Kirchhoff's Voltage Law* or the principle of *conservation of voltage.*

Remember the rule you've already learned about series dc circuits: The sum of the voltages across all the individual resistances adds up to the supply voltage. This statement is true as far as it goes, but it is an oversimplification, because it ignores polarity. The *polarity* of the potential difference across each resistance is *opposite* to the polarity of the potential difference across the battery. So when you add up the potential differences all the way around the circuit, taking polarity into account for every single component, you always get a net voltage of zero.

An example of Kirchhoff's Second Law is shown in Fig. 5-6. The voltage of the battery, E, has polarity opposite to the sum of the potential differences across the resistors, $E_1 + E_2 + E_3 + E_4$. Therefore, $E + E_1 + E_2 + E_3 + E_4 = 0$.



5-6 Kirchhoff's Second Law. The sum of the voltages across the resistances is equal to, but has opposite polarity from, the supply voltage. Therefore, $E + E_1 + E_2 + E_3 + E_4 = 0$. Illustration for Quiz Questions 15 and 16.

Problem 5-13

Refer to the diagram of Fig. 5-6. Suppose the four resistors have values of 50 Ω, 60 Ω, 70Ω, and 80 Ω, and that the current through each of them is 500 mA. What is the battery voltage, E?

Find the voltages E_1, E_2, E_3, and E_4 across each of the resistors. This can be done using Ohm's Law. For E_1, say with the 50-Ω resistor, calculate $E_1 = 0.500 \times 50 = 25$ V. In the same way, you can calculate $E_2 = 30$ V, $E_3 = 35$ V, and $E_4 = 40$ V. The supply voltage is the sum $E_1 + E_2 + E_3 + E_4 = 25 + 30 + 35 + 40 = 130$ V. Kirchhoff's Second Law tells us that the polarities of the voltages across the resistors are in the opposite direction from that of the battery.

Problem 5-14

In the situation shown by Fig. 5-6, suppose the battery provides 20 V. Suppose the resistors labeled with voltages E_1, E_2, E_3, and E_4 have ohmic values in the ratio 1:2:3:4 respectively. What is the voltage E_3?

This problem does not provide any information about current in the circuit, nor does it give you the exact resistances. But you don't need to know these things to solve for E_3. Regardless of what the actual ohmic values are, the ratio E_1:E_2:E_3:E_4 will be the same as long as the resistances are in the ratio 1:2:3:4. We can plug in any ohmic values we want for the values of the resistors, as long as they are in that ratio.

Let R_n be the resistance across which the voltage is E_n, where n can range from 1 to 4. Now that we have given the resistances specific names, suppose $R_1 = 1.0\ \Omega$, $R_2 = 2.0\ \Omega$, $R_3 = 3.0\ \Omega$, and $R_4 = 4.0\ \Omega$. These are in the proper ratio. The total resistance is $R = R_1 + R_2 + R_3 + R_4 = 1.0 + 2.0 + 3.0 + 4.0 = 10\ \Omega$. You can calculate the current as $I = E/R = 20/10 = 2.0$ A. Then the voltage E_3, across the resistance R_3, is given by Ohm's Law as $E_3 = IR_3 = 2.0 \times 3.0 = 6.0$ V.

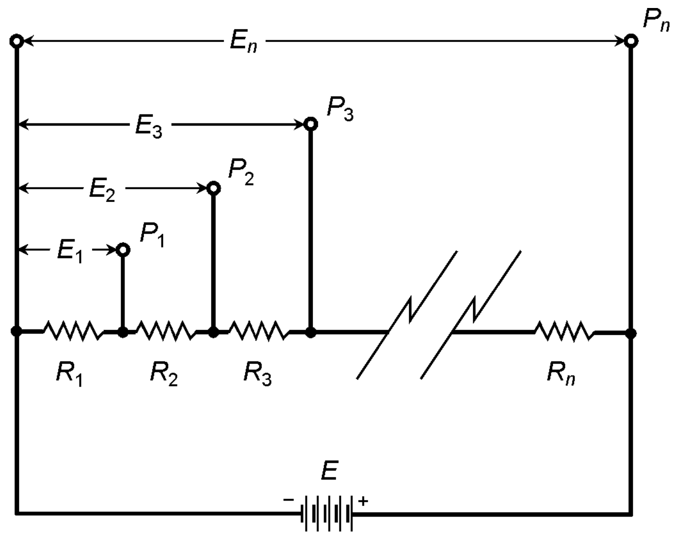
Voltage Divider Networks

Resistances in series produce ratios of voltages, and these ratios can be tailored to meet certain needs by means of *voltage divider networks.*

When a voltage divider network is designed and assembled, the resistance values should be as small as possible without causing too much current drain on the battery or power supply. (In practice, the optimum values depend on the nature of the circuit being designed. This is a matter for engineers, and specific details are beyond the scope of this course.) The reason for choosing the smallest possible resistances is that, when the divider is used with a circuit, you do not want that circuit to upset the operation of the divider. The voltage divider "fixes" the intermediate voltages best when the resistance values are as small as the current-delivering capability of the power supply will allow.

Figure 5-7 illustrates the principle of voltage division. The individual resistances are R_1, R_2, R_3, . . . , R_n. The total resistance is $R = R_1 + R_2 + R_3 + \ldots + R_n$. The supply voltage is E, and the current in the circuit is therefore $I = E/R$. At the various points P_1, P_2, P_3, . . . , P_n, the potential differences relative to the negative battery terminal are E_1, E_2, E_3, . . . , E_n, respectively. The last voltage, E_n, is the same as the battery voltage, E. All the other voltages are less than E, and ascend in succession, so that $E_1 < E_2 < E_3 < \ldots < E_n$. (The mathematical symbol < means "is less than.")

The voltages at the various points increase according to the sum total of the resistances up to each point, in proportion to the total resistance, multiplied by the supply voltage. Thus, the voltage E_1 is equal to ER_1/R. The voltage E_2 is equal to $E(R_1 + R_2)/R$. The voltage E_3 is equal to $E(R_1 + R_2 + R_3)/\mathrm{R}$. This process goes on for each of the voltages at points all the way up to $E_n = E(R_1 + R_2 + R_3 + \ldots + R_n)/R = ER/R = E$.

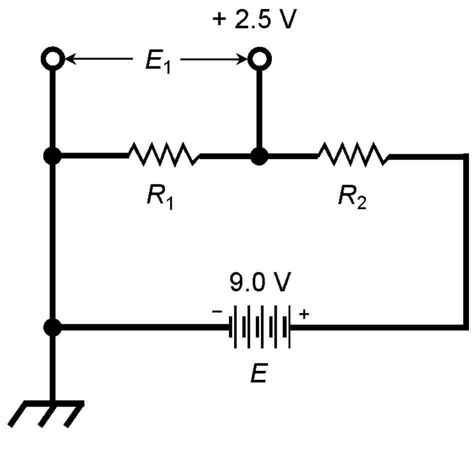


5-7 General arrangement for a voltage-divider circuit. Illustration for Quiz Questions 19 and 20.

Problem 5-15

Suppose you are building an electronic circuit, and the battery supplies 9.0 V. The minus terminal is at common (chassis) ground. You need to provide a circuit point where the dc voltage is +2.5 V. Give an example of a pair of resistors that can be connected in a voltage divider configuration, such that +2.5 V appears at some point.

Examine the schematic diagram of Fig. 5-8. There are infinitely many different combinations of resistances that will work here! Pick some total value, say $R = R_1 + R_2 = 1000\ \Omega$. Keep in mind that the ratio $R_1{:}R$ will always be the same as the ratio $E_1{:}E$. In this case, $E_1 = 2.5$ V, so $E_1{:}E = 2.5/9.0 = 0.28$. This means that you want the ratio $R_1{:}R$ to be equal to 0.28. You have chosen to



5-8 A voltage divider network in which 2.5 V dc is derived from a 9.0-V dc source.

make R equal to 1000 Ω. This means R_1 must be 280 Ω in order to get the ratio R_1:R = 0.28. The value of R_2 is the difference between R and R_1. That is 1000 − 280 = 720 Ω.

In a practical circuit, you would want to choose the smallest possible value for R. This might be less than 1000 Ω, or it might be more, depending on the nature of the circuit and the current-delivering capability of the battery. It's not the actual values of R_1 and R_2 that determine the voltage you get at the intermediate point, but their ratio.

Problem 5-16

What is the current I, in milliamperes, drawn by the entire network of series resistances in the situation described in Problem 5-15 and its solution?

Use Ohm's Law to get $I = E/R = 9.0/1000 = 0.0090$ A = 9.0 mA.

Problem 5-17

Suppose that it is all right for the voltage divider network to draw up to 100 mA of current in the situation shown by Fig. 5-8 and posed by Problem 5-15. You want to design the network to draw this amount of current, because that will offer the best voltage regulation for the circuit to be operated from the network. What values of resistances R_1 and R_2 should you use?

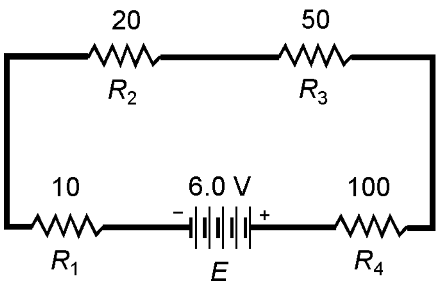
Calculate the total resistance first, using Ohm's Law. Remember to convert 100 mA to amperes! That means you use the figure I = 0.100 A in your calculations. Then $R = E/I = 9.0/0.100 = 90$ Ω. The ratio of resistances that you need is R_1:R_2 = 2.5/9.0 = 0.28. You should use R_1 = 0.28 × 90 = 25 Ω. The value of R_2 is the difference between R and R_1. That is, $R_2 = R - R_1 = 90 - 25 = 65$ Ω.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct answers. The answers are in the back of the book.

1. In a series-connected string of ornament bulbs, if one bulb gets shorted out, which of the following will occur?
   (a) All the other bulbs will go out.
   (b) The current in the string will go up.
   (c) The current in the string will go down.
   (d) The current in the string will stay the same.

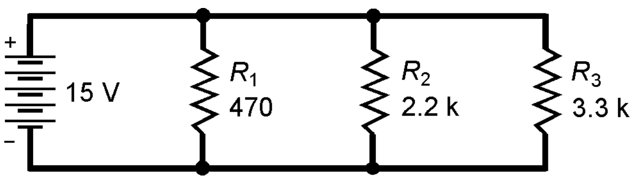
2. Imagine that four resistors are connected in series across a 6.0-V battery, and the ohmic values are R_1 = 10 Ω, R_2 = 20 Ω, R_3 = 50 Ω, and R_4 = 100 Ω, as shown in Fig. 5-9. What is the voltage across the resistance R_2?
   (a) 0.18 V
   (b) 33 mV
   (c) 5.6 mV
   (d) 0.67 V



5-9 Illustration for Quiz Questions 2, 3, 8, and 9. Resistance values are in ohms.

3. In the scenario shown by Fig. 5-9, what is the voltage across the combination of R_3 and R_4?
   (a) 0.22 V
   (b) 0.22 mV
   (c) 5.0 V
   (d) 3.3 V

4. Suppose three resistors are connected in parallel across a battery that delivers 15 V, and the ohmic values are $R_1 = 470\ \Omega$, $R_2 = 2.2\ \text{k}\Omega$, and $R_3 = 3.3\ \text{k}\Omega$, as shown in Fig. 5-10. The voltage across the resistance R_2 is
   (a) 4.4 V.
   (b) 5.0 V.
   (c) 15 V.
   (d) not determinable from the data given.



5-10 Illustration for Quiz Questions 4, 5, 6, 7, 10, and 11. Resistances are in ohms, where k indicates multiplication by 1000.

5. In the situation shown by Fig. 5-10, what is the current through R_2?
   (a) 6.8 mA
   (b) 43 mA
   (c) 0.15 A
   (d) 6.8 A

6. In the situation shown by Fig. 5-10, what is the total current drawn from the source?
   (a) 6.8 mA
   (b) 43 mA
   (c) 0.15 A
   (d) 6.8 A

7. In the situation shown by Fig. 5-10, suppose that resistor R_2 opens up. The current through the other two resistors will
   (a) increase.
   (b) decrease.
   (c) drop to zero.
   (d) not change.
8. Suppose that four resistors are connected in series with a 6.0-V supply, with values shown in Fig. 5-9. What is the power dissipated by the whole combination?
   (a) 0.2 W
   (b) 6.5 mW
   (c) 200 W
   (d) 6.5 W
9. In the situation shown by Fig. 5-9, what is the power dissipated by R_4?
   (a) 11 mW
   (b) 0.11 W
   (c) 0.2 W
   (d) 6.5 mW
10. Suppose that three resistors are in parallel as shown in Fig. 5-10. What is the power dissipated by the whole set of resistors?
   (a) 5.4 W
   (b) 5.4 μW
   (c) 650 W
   (d) 0.65 W
11. In the situation shown by Fig. 5-10, what is the power dissipated in resistance R_1?
   (a) 32 mW
   (b) 0.48 W
   (c) 2.1 W
   (d) 31 W
12. Fill in the blank in the following sentence to make it true: "In a series or parallel dc circuit, the sum of the ____s in each component is equal to the total ____ provided by the power supply."
   (a) current
   (b) voltage
   (c) wattage
   (d) resistance
13. Look at Fig. 5-5A. Suppose the resistors each have values of 33 Ω and the battery supplies 24 V. What is the current I_1?
   (a) 1.1 A
   (b) 0.73 A

(c) 0.36 A
(d) Not determinable from the information given

14. Look at Fig. 5-5B. Let each resistor have a value of 820 Ω. Suppose the top three resistors all lead to identical light bulbs. If I_1 = 50 mA and I_2 = 70 mA, what is the power dissipated in the resistor carrying current I_4?
(a) 33 W
(b) 40 mW
(c) 1.3 W
(d) It can't be found using the information given.

15. Refer to Fig. 5-6. Suppose the resistances R_1, R_2, R_3, and R_4 are in exactly the ratio 1:2:4:8 from left to right, and the battery supplies 30 V. What is the voltage E_2?
(a) 4.0 V
(b) 8.0 V
(c) 16 V
(d) It is not determinable from the data given.

16. Refer to Fig. 5-6. Suppose the resistances are each 3.3 kΩ, and the battery supplies 12 V. If the plus terminal of a dc voltmeter is placed between resistances R_1 and R_2 (with voltages E_1 and E_2 across them, respectively), and the minus terminal of the voltmeter is placed between resistances R_3 and R_4 (with voltages E_3 and E_4 across them, respectively), what will the meter register?
(a) 0.0 V
(b) 3.0 V
(c) 6.0 V
(d) 12 V

17. In a voltage divider network, the total resistance
(a) should be large to minimize current drain.
(b) should be as small as the power supply will allow.
(c) is not important.
(d) should be such that the current is kept to 100 mA.

18. The maximum voltage output from a voltage divider
(a) is a fraction of the power supply voltage.
(b) depends on the total resistance.
(c) is equal to the supply voltage.
(d) depends on the ratio of resistances.

19. Refer to Fig. 5-7. Suppose the battery voltage E is 18.0 V, and there are four resistances in the network such that R_1 = 100 Ω, R_2 = 22.0 Ω, R_3 = 33.0 Ω, and R_4 = 47.0 Ω. What is the voltage E_3 at P_3?
(a) 4.19 V
(b) 13.8 V

(c) 1.61 V

(d) 2.94 V

20. Refer to Fig. 5-7. Suppose the battery voltage is 12 V, and you want to obtain intermediate voltages of 3.0 V, 6.0 V, and 9.0 V. Suppose that a maximum of 200 mA is allowed to be drawn from the battery. What should the resistances, R_1, R_2, R_3, and R_4 be, respectively?

(a) 15 Ω, 30 Ω, 45 Ω, and 60 Ω

(b) 60 Ω, 45 Ω, 30 Ω, and 15 Ω

(c) 15 Ω, 15 Ω, 15 Ω, and 15 Ω

(d) There isn't enough information given here to design the circuit.

6
CHAPTER

Resistors

ALL ELECTRICAL COMPONENTS, DEVICES, AND SYSTEMS HAVE SOME RESISTANCE. IN EVERYDAY PRACTICE, there is no such thing as a perfect electrical conductor. You've seen some examples of circuits containing components that are deliberately designed to oppose the flow of current. These components are *resistors.* In this chapter, you'll learn all about them.

Purpose of the Resistor

Resistors play diverse roles in electrical and electronic equipment. Here are a few of the more common ways they are used.

Voltage Division

You've learned how voltage dividers can be designed using resistors. The resistors dissipate some power in doing this job, but the resulting voltages can provide the proper *biasing* of electronic circuits. This ensures, for example, that an amplifier or oscillator will function in the most efficient, reliable way possible.

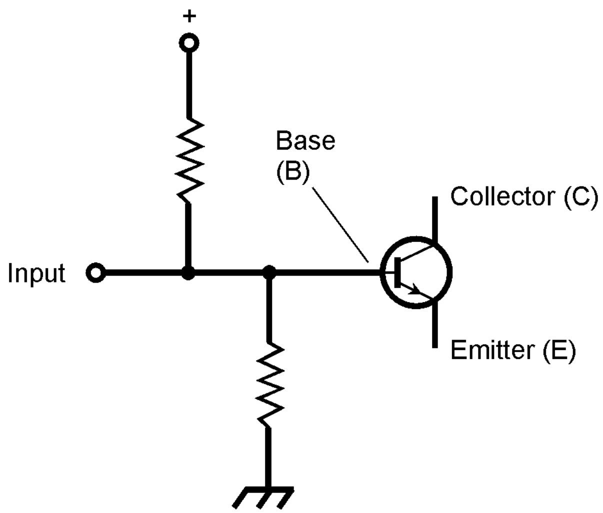
Bias

The term *bias* means, in the case of a *bipolar transistor*, a *field-effect transistor*, or a *vacuum tube*, that the control electrode—the *base*, *gate*, or *grid*—is provided with a certain voltage, or made to carry a certain current, relative to the *emitter*, *source*, or *cathode.* Networks of resistors can accomplish this.

A radio transmitting amplifier is biased differently than an oscillator or a low-level receiving amplifier. Sometimes voltage division is required for biasing. Other times it isn't necessary. Figure 6-1 shows a bipolar transistor whose base is biased using a pair of resistors in a voltage divider configuration.

Current Limiting

Resistors interfere with the flow of electrons in a circuit. Sometimes this is essential to prevent damage to a component or circuit. A good example is a receiving amplifier. A resistor can keep the

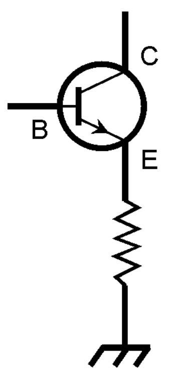


6-1 A pair of resistors can act as a voltage divider to bias the base of a transistor.

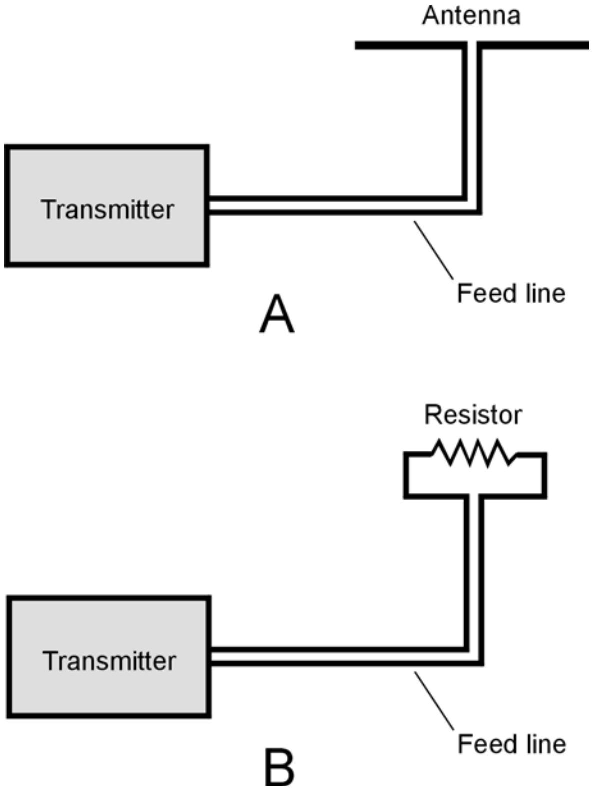
transistor from using up a lot of power just getting hot. Without resistors to limit or control the current, the transistor can be overstressed carrying direct current that doesn't contribute to the signal. Figure 6-2 shows a current-limiting resistor between the emitter of a bipolar transistor and electrical ground.

Power Dissipation

The dissipation of power in the form of heat is not always a bad thing. Sometimes a resistor can be used as a dummy component, so a circuit sees the resistor as if it were something more complicated. When testing a radio transmitter, for example, a resistor can be used to take the place of an antenna. This keeps the transmitter from interfering with communications on the airwaves. The transmitter output heats the resistor without radiating any signal. But as far as the transmitter knows, it's connected to a real antenna (Fig. 6-3)—and a perfect one, too, if the resistor has just the right ohmic value!



6-2 A resistor can limit the current that passes through the emitter of a transistor.

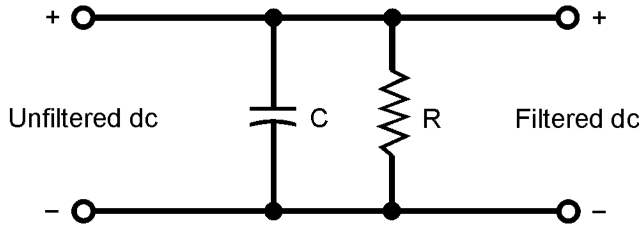


6-3 At A, a radio transmitter is connected to a real antenna. At B, the same transmitter is connected to a resistive dummy antenna.

Another situation in which power dissipation is useful is at the input of a power amplifier, such as the sort used in high-fidelity audio equipment. Sometimes the circuit *driving* the amplifier (supplying its input signal) has too much power. A resistor, or network of resistors, can dissipate this excess so that the amplifier doesn't get too much *drive.* In any type of amplifier, *overdrive* (an excessively strong input signal) can cause distortion, inefficiency, and other problems.

Bleeding Off Charge

In a high-voltage, dc power supply, capacitors are used to smooth out the fluctuations in the output. These capacitors acquire an electric charge, and they store it for a while. In some power supplies, these *filter capacitors* hold the full output voltage of the supply, say something like 750 V, even after the supply has been turned off, and even after it is unplugged from the wall outlet. If you attempt to repair such a power supply, you can be electrocuted by this voltage. *Bleeder resistors*, connected across the filter capacitors, drain their stored charge so that servicing the supply is not dangerous. In Fig. 6-4, the bleeder resistor, R, should have a value high enough so that it doesn't interfere with the operation of the power supply, but low enough so it will discharge the capacitor, C, in a short time after the power supply has been shut down.



6-4 A bleeder resistor (R) is connected across the filter capacitor (C) in a power supply.

It's a good idea to short out all filter capacitors, using a screwdriver with an insulated handle and wearing heavy, insulated gloves, before working on a dc power supply. Even if the supply has bleeder resistors, they might take a while to get rid of the residual charge. In addition, bleeder resistors can, and sometimes do, fail.

Impedance Matching

A more sophisticated application for resistors is in the *coupling* in a chain of amplifiers, or in the input and output circuits of amplifiers. In order to produce the greatest possible amplification, the *impedances* must agree between the output of a given amplifier and the input of the next. The same is true between a source of signal and the input of an amplifier. Also, this applies between the output of an amplifier and a *load*, whether that load is a speaker, a headset, or whatever.

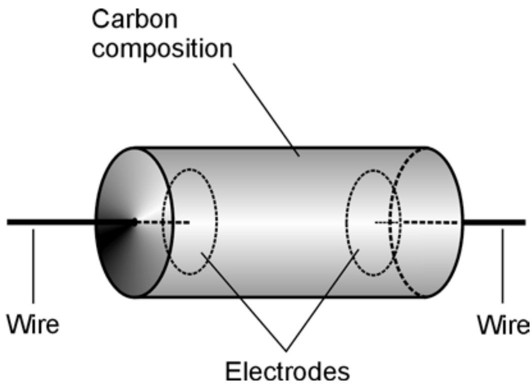
Impedance is the ac "big brother" of dc resistance. You will learn about impedance in Part 2 of this book.

Fixed Resistors

There are several ways in which *fixed resistors* (units whose resistance does not change, or cannot be adjusted) are manufactured. Here are the most common types.

Carbon-Composition Resistors

The cheapest method of making a resistor is to mix up powdered carbon (a fair electrical conductor) with some nonconductive substance, press the resulting claylike stuff into a cylindrical shape, and insert wire leads in the ends (Fig. 6-5). The resistance of the final product depends on the ratio



6-5 Construction of a carbon-composition resistor.

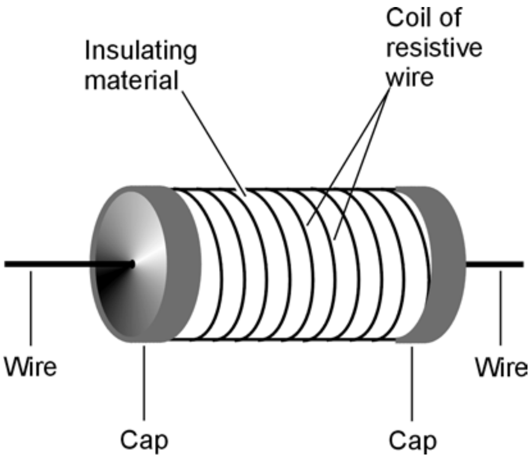
of carbon to the nonconducting material, and also on the physical distance between the wire leads. This results in a *carbon-composition resistor.*

Carbon-composition resistors can be manufactured in a wide range of resistance values. This kind of resistor also has the advantage of being *nonreactive*, meaning that it introduces almost pure resistance into the circuit, and not much capacitance or inductance. This makes carbon-composition resistors useful in radio receivers and transmitters.

Carbon-composition resistors dissipate power according to how big, physically, they are. Most of the carbon-composition resistors you see in electronics stores can handle $\frac{1}{4}$ W or $\frac{1}{2}$ W. There are $\frac{1}{8}$-W units available for miniaturized, low-power circuitry, and 1- or 2-W units for circuits where some electrical ruggedness is needed. Occasionally you'll see a carbon-composition resistor with a much higher power rating, but these are rare.

Wirewound Resistors

Another way to get resistance is to use a length of wire that isn't a good conductor. The wire can be wound around a cylindrical form as a coil (Fig. 6-6). The resistance is determined by how well the wire metal conducts, by its diameter or *gauge*, and by its stretched-out length. This type of component is called a *wirewound resistor.*

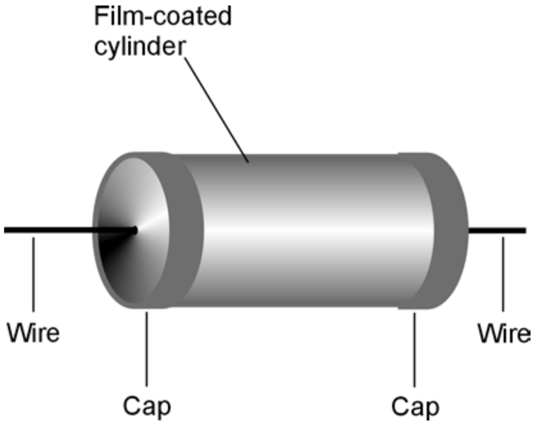


6-6 Construction of a wirewound resistor.

Wirewound resistors can be manufactured to have values within a very close range. They are *precision components.* Also, wirewound resistors can be made to handle large amounts of power. A disadvantage of wirewound resistors, in some applications, is that they act like inductors. This makes them unsuitable for use in most radio-frequency circuits. Wirewound resistors usually have low to moderate values of resistance.

Film-Type Resistors

Carbon, resistive wire, or some mixture of ceramic and metal can be applied to a cylindrical form as a film, or thin layer, in order to obtain a specific resistance. This type of component is called a *carbon-film resistor* or *metal-film resistor.* Superficially, it looks like a carbon-composition resistor, but the construction is different (Fig. 6-7).



6-7 Construction of a film-type resistor.

The cylindrical form is made of an insulating substance, such as porcelain. The film is deposited on this form by various methods, and the value tailored as desired. Metal-film resistors can be made to have nearly exact values. Film-type resistors usually have low to medium-high resistance.

A major advantage of film-type resistors is that they, like carbon-composition resistors, do not have much inductance or capacitance. A disadvantage, in some applications, is that they can't handle as much power as carbon-composition or wirewound types.

Integrated-Circuit (IC) Resistors

Resistors can be fabricated on a semiconductor wafer known as an *integrated circuit* (IC), also called a *chip*. The thickness, and the types and concentrations of impurities added, control the resistance of the component. Integrated-circuit resistors can handle only a tiny amount of power because of their small size.

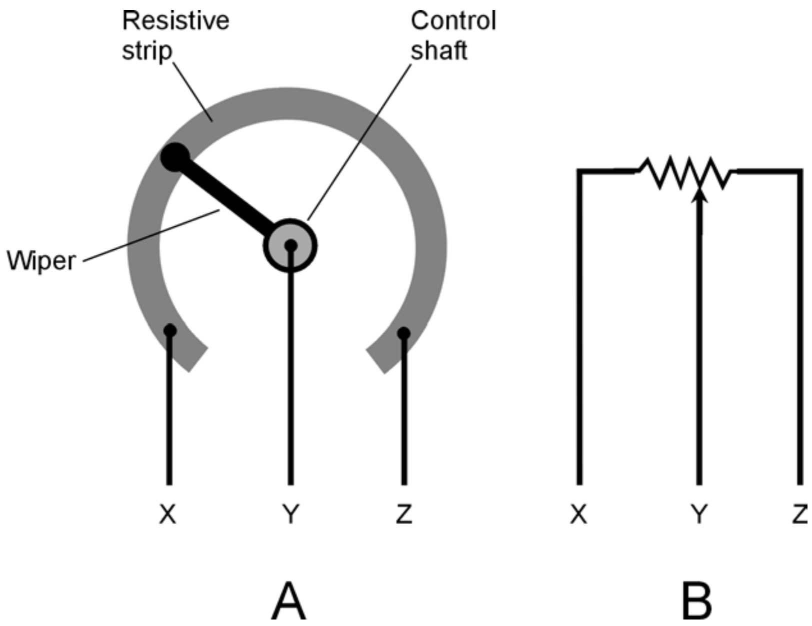
The Potentiometer

Figure 6-8 is a simplified drawing of the construction of a *potentiometer*, or variable resistor. A resistive strip, similar to that found on film-type fixed resistors, is bent into a nearly complete circle, and terminals are connected to either end. This forms a fixed resistance. To obtain the variable resistance, a sliding contact is attached to a rotatable shaft and bearing, and is connected to a third terminal. The resistance between the middle terminal and either of the end terminals can vary from zero up to the resistance of the whole strip.

Some potentiometers use a straight strip of resistive material, and the control moves up and down or from side to side. This type of variable resistor, called a *slide potentiometer*, is used in hi-fi audio *graphic equalizers*, as the volume controls in some hi-fi audio amplifiers, and in other applications when a linear scale is preferable to a circular scale. Potentiometers are manufactured to handle low levels of current, at low voltage.

Linear-Taper Potentiometer

One type of potentiometer uses a strip of resistive material whose density is constant all the way around. This results in a *linear taper*. The resistance between the center terminal and either end terminal changes at a steady rate as the control shaft is turned.



6-8 A simplified functional drawing of a rotary potentiometer (A), and the schematic symbol (B).

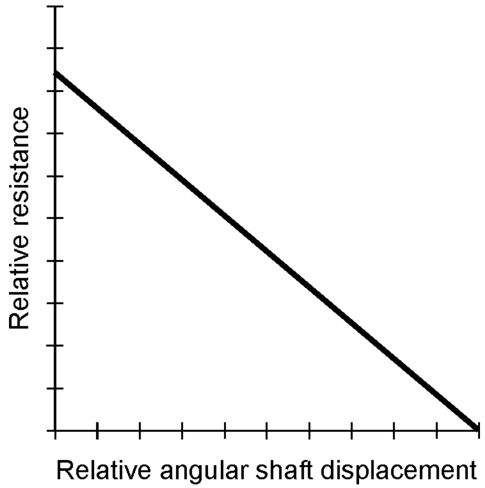
Suppose a *linear-taper potentiometer* has a value of zero to 280 Ω. In most units the shaft can be rotated through about 280°, or a little more than three-quarters of a circle. The resistance between the center and one end terminal will increase right along with the number of angular degrees that the shaft is turned. The resistance between the center and the other end terminal will be equal to 280 minus the number of degrees the shaft is turned. The resistance is a *linear function* of the angular shaft position.

Linear-taper potentiometers are commonly used in electronic test instruments and in various consumer electronic devices. Figure 6-9 is a graph of relative resistance versus relative angular shaft displacement for a linear-taper potentiometer.

Audio-Taper Potentiometer

In some applications, linear taper potentiometers don't work well. The volume control of a radio receiver or hi-fi audio amplifier is a good example. Humans perceive sound intensity according to the *logarithm* of the actual sound power. If you use a linear-taper potentiometer as the volume control for a radio or other sound system, the sound volume will vary too slowly in some parts of the control range, and too fast in other parts of the control range.

To compensate for the way in which people perceive sound level, an *audio-taper potentiometer* is used. In this device, the resistance between the center and end terminal increases as a nonlinear function of the angular shaft position. The device is sometimes called a *logarithmic-taper potentiometer* or *log-taper potentiometer* because the nonlinear function is logarithmic. This precisely compensates for the way the human ear-and-brain "machine" responds to sounds of variable intensity. Audio-taper potentiometers are manufactured so that as you turn the shaft, the sound intensity



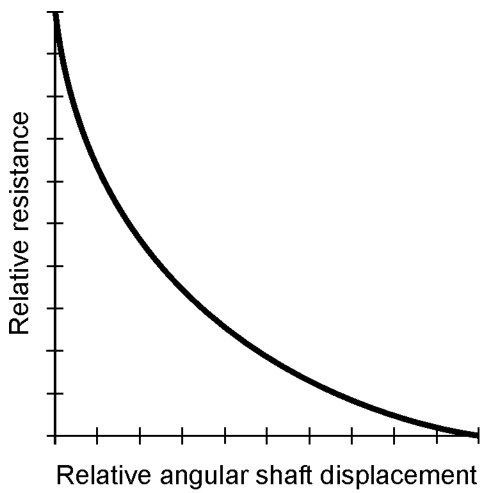
6-9 Resistance as a function of angular displacement for a linear-taper potentiometer.

seems to increase in a smooth, natural way. Figure 6-10 is a graph of relative resistance versus relative angular shaft displacement for an audio-taper potentiometer.

The Rheostat

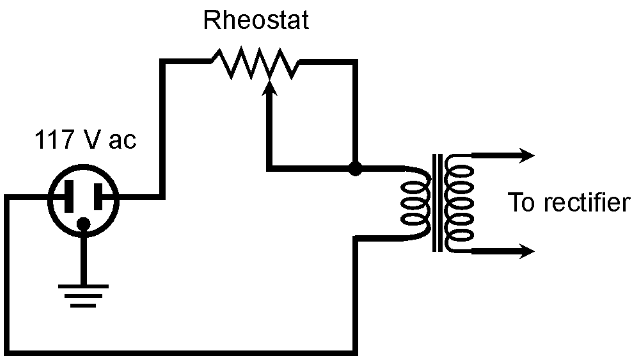
A variable resistor can be made from a wirewound element, rather than a solid strip of material. This is called a *rheostat.* It can have either a rotary control or a sliding control. This depends on whether the resistive wire is wound around a donut-shaped form (*toroid*) or a cylindrical form (*solenoid*). Rheostats have inductance as well as resistance. They share the advantages and disadvantages of fixed wirewound resistors.

A rheostat is not continuously adjustable, as is a potentiometer. This is because the movable contact slides along from turn to turn of the wire coil. The smallest possible increment is the resistance in one turn of the coil.



6-10 Resistance as a function of angular displacement for an audio-taper potentiometer.

6-11 Connection of a rheostat in a variable-voltage power supply.



Rheostats are used in high-voltage, high-power applications. A good example is in a variable-voltage power supply. This kind of supply uses a transformer that steps up the voltage from the 117-V utility mains, and diodes to change the ac to dc. The rheostat can be placed between the utility outlet and the transformer (Fig. 6-11). This results in a variable voltage at the power-supply output.

The Decibel

As stated in the preceding paragraphs, perceived levels of sound change according to the logarithm of the actual sound power level. The same is true for various other phenomena, too, such as visible-light intensity and radio-frequency signal strength. Specialized units have been defined to take this into account.

The fundamental unit of sound-level change is called the *decibel*, symbolized as dB. A change of +1 dB is the minimum increase in sound level that you can detect if you are expecting it. A change of −1 dB is the minimum detectable decrease in sound volume, when you are anticipating the change. Increases in volume are given positive decibel values, and decreases in volume are given negative decibel values.

If you aren't expecting the level of sound to change, then it takes about +3 dB or −3 dB to make a noticeable difference.

Changes in intensity, when expressed in decibels, are sometimes called *gain* and *loss.* Positive decibel changes represent gain, and negative decibel changes represent loss. The sign (plus or minus) is usually absent when speaking of changes in terms of decibel gain or decibel loss. If you say that a certain system causes 5 dB of loss, you are saying that the gain of that circuit is −5 dB.

Calculating Decibel Values

Decibel values are calculated according to the logarithm of the ratio of change. Suppose a sound produces a power of P watts on your eardrums, and then it changes (either getting louder or softer) to a level of Q watts. The change in decibels is obtained by dividing out the ratio Q/P, taking its base-10 logarithm (symbolized as $\log_{10}$ or simply as log), and then multiplying the result by 10. Mathematically:

$$\text{dB} = 10 \log (Q/P)$$

As an example, suppose a speaker emits 1 W of sound, and then you turn up the volume so that it emits 2 W of sound power. Then $P = 1$ and $Q = 2$, and dB = 10 log (2/1) = 10 log 2 = 10 × 0.3 = 3 dB. This is the minimum detectable level of volume change if you aren't expecting it: doubling of the actual sound power!

If you turn the volume level back down again, then $P/Q = 1/2 = 0.5$, and you can calculate dB = 10 log 0.5 = 10 × −0.3 = −3 dB.

A gain or loss of 10 dB (that is, a change of +10 dB or −10 dB, often shortened to ±10 dB) represents a 10-fold increase or decrease in sound power. A change of ±20 dB represents a 100-fold increase or decrease in sound power. It is not unusual to encounter sounds that vary in intensity over ranges of ±60 dB, which represents a 1,000,000-fold increase or decrease in sound power!

Sound Power in Terms of Decibels

The preceding formula can be worked inside out, so that you can determine the final sound power, given the initial sound power and the decibel change. To do this, you use the *inverse of the logarithmic function*, symbolized as $\log^{-1}$ or antilog. This function, like the logarithmic function, can be performed by any good scientific calculator, or by the calculator program in a personal computer when set to scientific mode.

Suppose the initial sound power is P, and the change in decibels is dB. Let Q be the final sound power. Then:

$$Q = P \text{ antilog } (\text{dB}/10)$$

As an example, suppose the initial power, P, is 10 W, and the perceived volume change is −3 dB. Then the final power, Q, is equal to 10 antilog (−3/10) = 10 × 0.5 = 5 W.

Decibels in the Real World

Sound levels are sometimes specified in decibels relative to the *threshold of hearing*, defined as the faintest possible sound that a person can detect in a quiet room, assuming his or her hearing is normal. This threshold is assigned the value 0 dB. Other sound levels can then be quantified as figures such as 30 dB or 75 dB.

If a certain noise has a loudness of 30 dB, that means it's 30 dB above the threshold of hearing, or 1000 times as loud as the quietest detectable noise. A noise at 60 dB is 1,000,000 (or 10^6) times as powerful as a sound at the threshold of hearing. *Sound-level meters* are used to determine the decibel levels of various noises and acoustic environments.

A typical conversation occurs at a level of about 70 dB. This is 10,000,000 (or 10^7) times the threshold of hearing, in terms of actual sound power. The roar of the crowd at a rock concert might be 90 dB, or 1,000,000,000 (10^9) times the threshold of hearing. A sound at 100 dB, typical of the music at a large rock concert if you are sitting in the front row, is 10,000,000,000 (10^{10}) times as loud, in terms of power, as a sound at the threshold of hearing.

Resistor Specifications

When choosing a resistor for a particular application in an electrical or electronic device, it's important to get a unit that has the correct properties, or *specifications.* Here are some of the most important specifications to watch for.

Ohmic Value

In theory, a resistor can have any ohmic value from the lowest possible (such as a shaft of solid silver) to the highest (dry air). In practice, it is unusual to find resistors with values less than about 0.1 Ω or more than about 100 MΩ.

Resistors are manufactured with ohmic values in power-of-10 multiples of 1.0, 1.2, 1.5, 1.8, 2.2, 2.7, 3.3, 3.9, 4.7, 5.6, 6.8, and 8.2. Thus, you will often see resistors with values of 47 Ω, 180 Ω, 6.8 kΩ, or 18 MΩ, but hardly ever with values such as 384 Ω, 4.54 kΩ, or 7.297 MΩ.

In addition to these standard values, there are others that are used for resistors made with greater precision, or tighter *tolerance*. These are power-of-10 multiples of 1.1, 1.3, 1.6, 2.0, 2.4, 3.0, 3.6, 4.3, 5.1, 6.2, 7.5, and 9.1.

Tolerance

The first set of numbers above represents standard resistance values available in tolerances of plus or minus 10 percent (±10%). This means that the resistance might be as much as 10 percent more or 10 percent less than the indicated amount. In the case of a 470-Ω resistor, for example, the value can be larger or smaller than the rated value by as much as 47 Ω, and still be within tolerance. That's a range of 423 to 517 Ω.

Tolerance is calculated according to the specified value of the resistor, not the actual value. You might measure the value of a 470-Ω resistor and find it to be 427 Ω, and it would be within ±10% of the specified value. But if it measures 420 Ω, it's outside the rated range, and is therefore a reject. The second set, along with the first set, of numbers represents standard resistance values available in tolerances of plus or minus 5 percent (±5%). A 470-Ω, 5 percent resistor will have an actual value of 470 Ω plus or minus 24 Ω, or a range of 446 to 494 Ω.

Some resistors are available in tolerances tighter than ±5%. These precision units are employed in circuits where a little error can make a big difference. In most audio and radio-frequency oscillators and amplifiers, the ±10% or ±5% tolerance is good enough. In many cases, even a ±20% tolerance is satisfactory.

Power Rating

All resistors are given a specification that determines how much power they can safely dissipate. Typical values are ¼ W, ½ W, and 1 W. Units also exist with ratings of ⅛ W or 2 W. These dissipation ratings are for *continuous duty*, meaning they can dissipate this amount of power constantly and indefinitely.

You can figure out how much current a given resistor can handle by using the formula for power (P) in terms of current (I) and resistance (R). That formula, you should recall, is $P = I^2R$. Work this formula backward, plugging in the power rating in watts for P and the resistance in ohms for R, and solve for the current I in amperes. Alternatively, you can find the square root of P/R.

The power rating for a given resistor can, in effect, be increased by using a network of 2 × 2, 3 × 3, 4 × 4, or more units in series-parallel. If you need a 47-Ω, 45-W resistor, but all you have is a bunch of 47-Ω, 1-W resistors, you can make a 7 × 7 network in series-parallel, and this will handle 49 W.

Resistor power dissipation ratings are specified with a margin for error. A good engineer never tries to take advantage of this and use, say, a ¼-W unit in a situation that needs to draw 0.27 W. In fact, good engineers usually include their own safety margin. Allowing 10 percent, a ¼-W resistor should not be called upon to handle more than about 0.225 W.

Temperature Compensation

All resistors change value when the temperature changes dramatically. And because resistors dissipate power, they can get hot just because of the current they carry. Often, this current is so tiny that it doesn't appreciably heat the resistor. But in some cases it does, and the resistance will change. Then a circuit might behave differently than it did when the resistor was still cool.

There are various ways to approach problems of resistors changing value when they get hot. One method is to use specially manufactured resistors that do not appreciably change value when they get hot. Such units are called *temperature-compensated.* But one of these can cost several times as much as an ordinary resistor. Another approach is to use a power rating that is much higher than the actual dissipated power in the resistor. This will keep the resistor from getting very hot. Still another scheme is to use a series-parallel network of identical resistors to increase the power dissipation rating. Alternatively, you can take several resistors, say three of them, each with about three times the intended resistance, and connect them all in parallel. Or you can take several resistors, say four of them, each with about one-fourth the intended resistance, and connect them in series.

It is unwise to combine resistors with different values. This can result in one of them taking most of the load while the others "loaf," and the combination will be no better than the single hot resistor you started with.

How about using two resistors with half (or twice) the value you need, but with *opposite* resistance-versus-temperature characteristics, and connecting them in series or parallel? It is tempting to suppose that if you do this, the component whose resistance decreases with heat (*negative temperature coefficient*) will have a canceling-out effect on the component whose resistance goes up (*positive temperature coefficient*). This can sometimes work, but in practice it's difficult to find a pair of resistances that will do this job just right.

The Color Code for Resistors

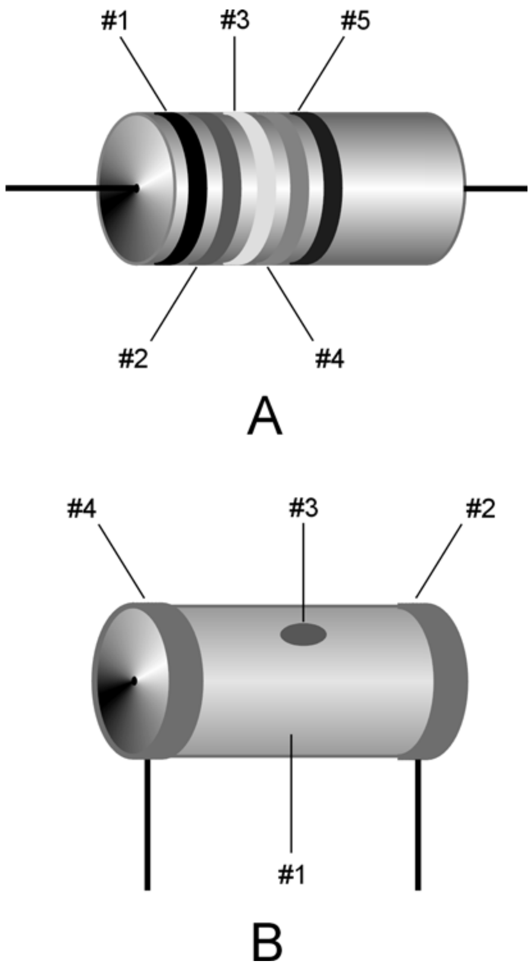
Some resistors have *color bands* that indicate their values and tolerances. You'll see three, four, or five bands around carbon-composition resistors and film resistors. Other units are large enough so that the values can be printed on them in ordinary numerals.

On resistors with *axial leads* (wires that come straight out of both ends), the first, second, third, fourth, and fifth bands are arranged as shown in Fig. 6-12A. On resistors with *radial leads* (wires that come off the ends at right angles to the axis of the component body), the colored regions are arranged as shown in Fig. 6-12B. The first two regions represent numbers 0 through 9, and the third region represents a multiplier of 10 to some power. (For the moment, don't worry about the fourth and fifth regions.) Refer to Table 6-1.

Suppose you find a resistor whose first three bands are yellow, violet, and red, in that order. Then the resistance is 4700 Ω. Read yellow = 4, violet = 7, red = $\times 100$. As another example, suppose you find a resistor with bands of blue, gray, orange. Refer to Table 6-1 and determine blue = 6, gray = 8, orange = $\times 1000$. Therefore, the value is 68,000 Ω = 68 kΩ.

The fourth band, if there is one, indicates tolerance. If it's silver, it means the resistor is rated at $\pm 10\%$. If it's gold, the resistor is rated at $\pm 5\%$. If there is no fourth band, the resistor is rated at $\pm 20\%$.

The fifth band, if there is one, indicates the maximum percentage that the resistance can be expected to change after 1000 hours of use. A brown band indicates a maximum change of $\pm 1\%$ of the rated value. A red band indicates $\pm 0.1\%$. An orange band indicates $\pm 0.01\%$. A yellow band indicates $\pm 0.001\%$. If there is no fifth band, it means that the resistor might deviate by more than $\pm 1\%$ of the rated value after 1000 hours of use.



6-12 At A, locations of color-code bands on a resistor with axial leads. At B, locations of color code designators on a resistor with radial leads.

Table 6-1. The color code for the first three bands that appear on fixed resistors. See text for discussion of the fourth and fifth bands.

Color of band	Numeral (first and second bands)	Multiplier (third band)
Black	0	1
Brown	1	10
Red	2	100
Orange	3	1000 (1 k)
Yellow	4	10^4 (10 k)
Green	5	10^5 (100 k)
Blue	6	10^6 (1 M)
Violet	7	10^7 (10 M)
Gray	8	10^8 (100 M)
White	9	10^9 (1000 M or 1 G)

A competent engineer or technician always tests a resistor with an ohmmeter before installing it in a circuit. If the component happens to be labeled wrong, or if it is defective, it's easy to catch this problem while assembling or servicing a circuit. But once the circuit is all together, and it won't work because some resistor is labeled wrong or is bad, it's difficult to troubleshoot.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct. Answers are in the back of the book.

1. Proper biasing in an amplifier circuit
   (a) causes it to oscillate.
   (b) prevents an impedance match.
   (c) can be obtained using a voltage divider network.
   (d) maximizes current flow.

2. A transistor can be protected from needless overheating by
   (a) a current-limiting resistor.
   (b) bleeder resistors.
   (c) maximizing the drive.
   (d) shorting out the power supply when the circuit is off.

3. A bleeder resistor
   (a) is connected across the capacitor in a power supply.
   (b) keeps a transistor from drawing too much current.
   (c) prevents an amplifier from being overdriven.
   (d) optimizes the efficiency of an amplifier.

4. Carbon-composition resistors
   (a) can handle gigantic levels of power.
   (b) have capacitance or inductance along with resistance.
   (c) have essentially no capacitance or inductance.
   (d) work better for ac than for dc.

5. A logical place for a wirewound resistor is
   (a) in a radio-frequency amplifier.
   (b) in a circuit where a noninductive resistor is called for.
   (c) in a low-power radio-frequency circuit.
   (d) in a high-power dc circuit.

6. A metal-film resistor
   (a) is made using a carbon-based paste.
   (b) does not have much inductance.

(c) can dissipate large amounts of power.

(d) has considerable inductance.

7. What type of resistor, or combination of resistors, would you use as the meter-sensitivity control in a test instrument, when continuous adjustment is desired?

(a) A set of switchable, fixed resistors

(b) A linear-taper potentiometer

(c) An audio-taper potentiometer

(d) A wirewound resistor

8. What type of resistor, or combination of resistors, would you use as the volume control in a stereo compact-disc (CD) player?

(a) A set of switchable, fixed resistors

(b) A linear-taper potentiometer

(c) An audio-taper potentiometer

(d) A wirewound resistor

9. If a sound triples in actual power level, approximately what is this, expressed in decibels?

(a) +3 dB

(b) +5 dB

(c) +6 dB

(d) +9 dB

10. Suppose a sound changes in volume by −13 dB. If the original sound power is 1.0 W, what is the final sound power?

(a) 13 W

(b) 77 mW

(c) 50 mW

(d) There is not enough information given here to answer this question.

11. The sound from a portable radio is at a level of 50 dB. How many times the threshold of hearing is this, in terms of actual sound power?

(a) 50

(b) 169

(c) 5000

(d) 100,000

12. An advantage of a rheostat over a potentiometer is the fact that

(a) a rheostat can handle higher frequencies.

(b) a rheostat is more precise.

(c) a rheostat can handle more current.

(d) a rheostat works better with dc.

13. A resistor is specified as having a value of 68 Ω, but is measured with an ohmmeter as 63 Ω. The value is off by which of the following percentages?

(a) 7.4%
(b) 7.9%
(c) 5%
(d) 10%

14. Suppose a resistor is rated at 3.3 kΩ ±5%. This means it can be expected to have a value between

(a) 2970 Ω and 3630 Ω.
(b) 3295 Ω and 3305 Ω.
(c) 3135 Ω and 3465 Ω.
(d) 2.8 kΩ and 3.8 kΩ.

15. A package of resistors is rated at 56 Ω ±10%. You test them with an ohmmeter. Which of the following values indicates a reject?

(a) 50.0 Ω
(b) 53.0 Ω
(c) 59.7 Ω
(d) 61.1 Ω

16. A resistor has a value of 680 Ω, and you expect that it will have to draw 1 mA maximum continuous current in a circuit you're building. What power rating is good for this application, but not needlessly high?

(a) ¼ W
(c) ½ W
(c) 1 W
(d) 2 W

17. Suppose a 1-kΩ resistor will dissipate 1.05 W, and you have a good supply of 1-W resistors of various ohmic values. If there's room for 20 percent resistance error, the cheapest solution is to use

(a) four 1-kΩ, 1-W resistors in series-parallel.
(b) a pair of 2.2-kΩ, 1-W resistors in parallel.
(c) a set of three 3.3-kΩ, 1-W resistors in parallel.
(d) a single 1-kΩ, 1-W resistor, because all manufacturers allow for a 10 percent margin of safety when rating resistors for their power-handling capability.

18. Suppose a carbon-composition resistor has the following colored bands on it: red, red, red, gold. This indicates a resistance of

(a) 22 Ω.
(b) 220 Ω.
(c) 2.2 kΩ.
(d) 22 kΩ.

19. The actual resistance of the component described in the previous question can be expected to vary above or below the specified ohmic value by up to what amount?

    (a) 11 Ω

    (b) 110 Ω

    (c) 22 Ω

    (d) 220 Ω

20. Suppose a carbon-composition resistor has the following colored bands on it: gray, red, yellow. This unit can be expected to have a value within approximately what range?

    (a) 660 kΩ to 980 kΩ

    (b) 740 kΩ to 900 kΩ

    (c) 7.4 kΩ to 9.0 kΩ

    (d) The manufacturer does not make any claim.

7
CHAPTER

Cells and Batteries

IN ELECTRICITY AND ELECTRONICS, A *CELL* IS A UNIT SOURCE OF DC ENERGY. WHEN TWO OR MORE cells are connected in series, the result is known as a *battery.* There are many types of cells and batteries, and new types are constantly being invented.

Electrochemical Energy

Early in the history of electrical science, laboratory physicists found that when metals came into contact with certain chemical solutions, voltages appeared between the pieces of metal. These were the first *electrochemical cells.*

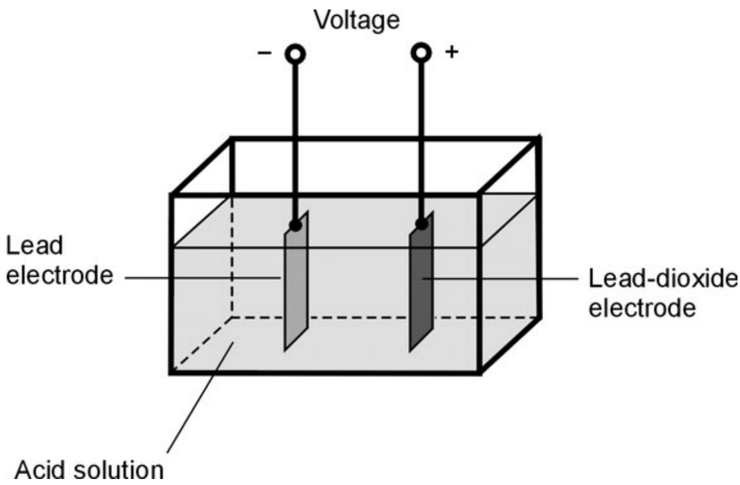
A piece of lead and a piece of lead dioxide immersed in an acid solution (Fig. 7-1) acquire a persistent potential difference. This can be detected by connecting a galvanometer between the pieces of metal. A resistor of about 1000 Ω must be used in series with the galvanometer in experiments of this kind, because connecting the galvanometer directly will cause too much current to flow, possibly damaging the galvanometer and causing the acid to boil.

The chemicals and the metal have an inherent ability to produce a constant exchange of charge carriers. If the galvanometer and resistor are left hooked up between the two pieces of metal for a long time, the current will gradually decrease, and the electrodes will become coated. All the *chemical energy* in the acid will have been turned into electrical energy as current in the wire and galvanometer. In turn, this current will have heated the resistor (another form of kinetic energy), and escaped into the air and into space.

Primary and Secondary Cells

Some electrical cells, once their chemical energy has all been changed to electricity and used up, must be thrown away. These are called *primary cells.* Other kinds of cells, such as the lead-and-acid type, can get their chemical energy back again by means of *recharging.* Such a cell is a *secondary cell.*

Primary cells include the ones you usually put in a flashlight, in a transistor radio, and in various other consumer devices. They use dry electrolyte pastes along with metal electrodes. They go by names such as *dry cell*, *zinc-carbon cell*, or *alkaline cell.* Go into a department store and find a rack of batteries, and you'll see various sizes and types of primary cells, such as *AAA batteries*, *D batteries*,



7-1 Simplified drawing of the construction of a lead-acid electrochemical cell.

camera batteries, and *watch batteries.* (These are actually cells, not true batteries.) You'll also see 9-V *transistor batteries* and large 6-V *lantern batteries.*

Secondary cells can also be found in consumer stores. *Nickel-based cells* are common. The most common sizes are AA, C, and D. These cost several times as much as ordinary dry cells, and a charging unit also costs a few dollars. But if you take care of them, these rechargeable cells can be used hundreds of times and will pay for themselves several times over if you use a lot of batteries in everyday life.

The battery in your car is made from secondary cells connected in series. These cells recharge from the alternator or from an outside charging unit. This battery has cells like the one in Fig. 7-1. It is dangerous to short-circuit the terminals of such a battery, because the acid (sulfuric acid) can bubble up and erupt out of the battery casing. Serious skin and eye injuries can result. In fact, it's a bad idea to short-circuit any cell or battery, because it can get extremely hot and cause a fire, or rupture and damage surrounding materials, wiring, and components.

The Weston Standard Cell

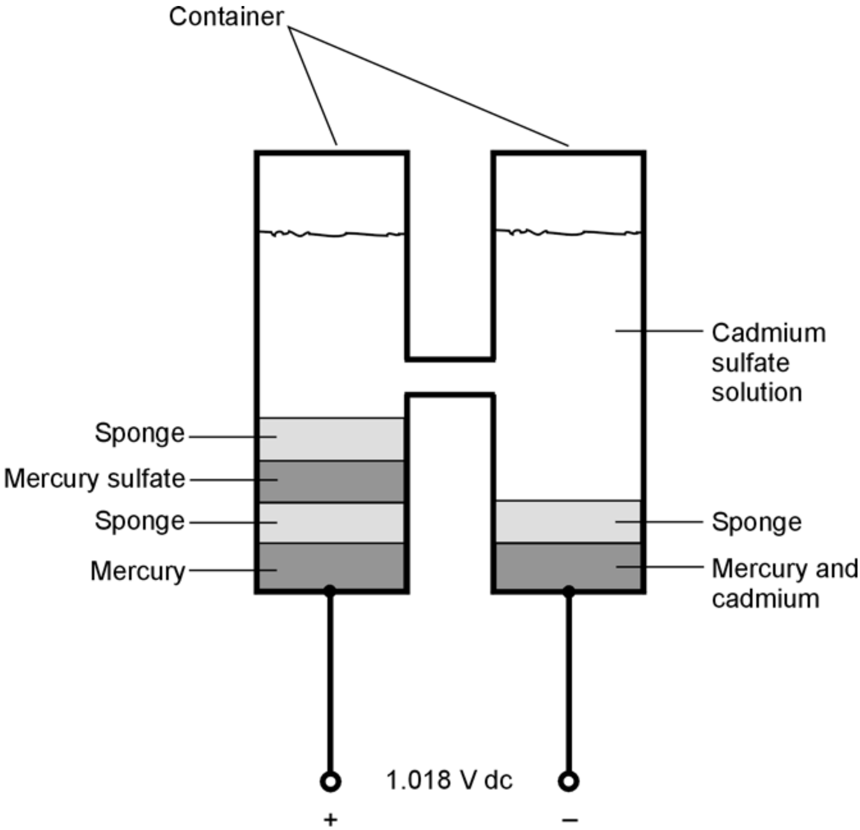
Most electrochemical cells produce 1.2 to 1.8 V. Different types vary slightly. A mercury cell has a voltage that is a little less than that of a zinc-carbon or alkaline cell. The voltage of a cell can also be affected by variables in the manufacturing process. Most consumer-type dry cells can be assumed to produce 1.5 V.

There are certain cells whose voltages are predictable and exact. These are called *standard cells.* A good example is the *Weston cell*, which produces 1.018 V at room temperature. It has a solution of cadmium sulfate, a positive electrode made from mercury sulfate, and a negative electrode made from mercury and cadmium. The device is set up in a container, as shown in Fig. 7-2.

Storage Capacity

Recall that the common electrical units of energy are the watt-hour (Wh) and the kilowatt-hour (kWh). Any electrochemical cell or battery has a certain amount of electrical energy that can be obtained from it, and this can be specified in watt-hours or kilowatt-hours. More often, though, it's given in *ampere-hours* (Ah).

A battery with a rating of 2 Ah can provide 2 A for 1 h, or 1 A for 2 h, or 100 mA for 20 h.

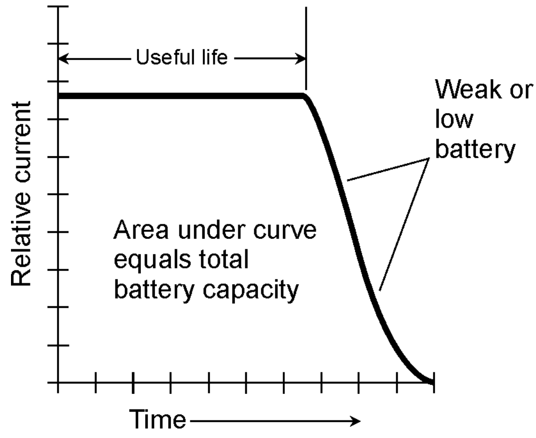


7-2 Simplified drawing of the construction of a Weston standard cell.

There are infinitely many possibilities here, as long as the product of the current in amperes and the use time in hours is equal to 2. The limitations are the *shelf life* at one extreme, and the *maximum deliverable current* at the other. Shelf life is the length of time the battery will last if it is never used; this can be years. The maximum deliverable current is the highest amount of current that the battery can provide before its voltage drops because of its own *internal resistance.*

Small cells have storage capacity of a few milliampere-hours (mAh) up to 100 or 200 mAh. Medium-sized cells can supply 500 mAh to 1 Ah. Large automotive or truck batteries can provide upward of 50 Ah. The energy capacity in watt-hours is the ampere-hour capacity multiplied by the battery voltage.

An *ideal cell* or *ideal battery* (a theoretically perfect cell or battery) delivers a constant current for a while, and then the current starts to drop (Fig. 7-3). Some types of cells and batteries approach this level of perfection, which is represented by a *flat discharge curve.* But many cells and batteries are far from perfect; they deliver current that declines gradually, almost right from the start. When the current that a battery can provide has tailed off to about half of its initial value, the cell or battery is said to be *weak.* At this time, it should be replaced. If it's allowed to run all the way out, until the current actually goes to zero, the cell or battery is *dead.* The area under the curve in Fig. 7-3 is a graphical representation the total capacity of the cell or battery in ampere-hours.



7-3 A flat discharge curve. This is considered ideal.

Grocery Store Cells and Batteries

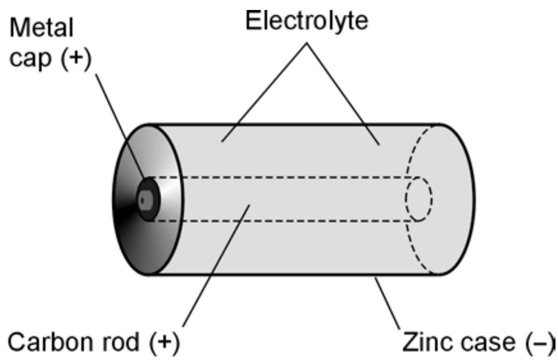
The cells you see in grocery stores, department stores, drugstores, and hardware stores provide 1.5 V, and are available in sizes known as AAA (very small), AA (small), C (medium large), and D (large). Batteries are widely available that deliver 6 or 9 V.

Zinc-Carbon Cells

Figure 7-4 is a translucent drawing of a *zinc-carbon cell.* The zinc forms the case and is the negative electrode. A carbon rod serves as the positive electrode. The electrolyte is a paste of manganese dioxide and carbon. Zinc-carbon cells are inexpensive and are good at moderate temperatures and in applications where the current drain is moderate to high. They are not very good in extreme cold.

Alkaline Cells

The *alkaline cell* has granular zinc as the negative electrode, potassium hydroxide as the electrolyte, and a device called a *polarizer* as the positive electrode. The construction is similar to that of the zinc-carbon cell. An alkaline cell can work at lower temperatures than a zinc-carbon cell. It lasts longer in most electronic devices, and is therefore preferred for use in transistor radios, calculators,



7-4 Simplified drawing of the construction of a zinc-carbon electrochemical cell.

and portable cassette players. Its shelf life is much longer than that of a zinc-carbon cell. As you might expect, it costs more.

Transistor Batteries

A *transistor battery* consists of six tiny zinc-carbon or alkaline cells in series. Each of the six cells supplies 1.5 V. Thus, the battery supplies 9 V. Even though these batteries have more voltage than individual cells, the total energy available from them is less than that from a C cell or D cell. This is because the electrical energy that can be obtained from a cell or battery is directly proportional to the amount of chemical energy stored in it, and this, in turn, is a direct function of the *volume* (physical size) of the cell or the *mass* (quantity of chemical matter) of the cell. Cells of size C or D have more volume and mass than a transistor battery, and therefore contain more stored energy for the same chemical composition.

Transistor batteries are used in low-current electronic devices such as remote-control garage-door openers, television (TV) and hi-fi remote controls, and electronic calculators.

Lantern Batteries

The *lantern battery* has much greater mass than a common dry cell or transistor battery, and consequently it lasts much longer and can deliver more current. Lantern batteries are usually rated at 6 V, and consist of four good-size zinc-carbon or alkaline cells. Two lantern batteries connected in series make a 12-V battery that can power a 5-W citizens band (CB) or ham radio transceiver for a while. They're also good for scanner radio receivers in portable locations, for camping lamps, and for other medium-power needs.

Miniature Cells and Batteries

In recent years, cells and batteries—especially cells—have become available in many different sizes and shapes besides the old cylindrical cells, transistor batteries, and lantern batteries. These are used in wristwatches, small cameras, and various microminiature electronic devices.

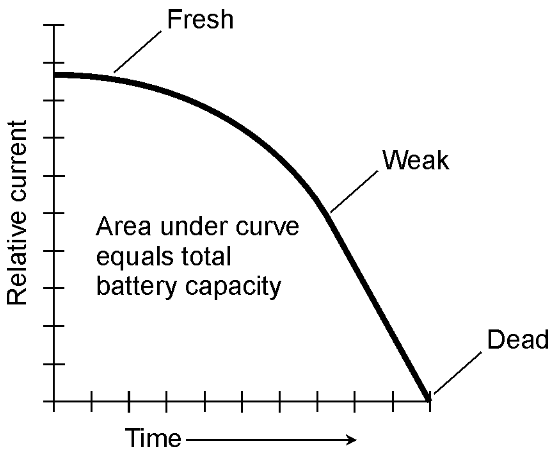
Silver-Oxide Cells and Batteries

A *silver-oxide cell* is usually found in a buttonlike shape, and can fit inside a small wristwatch. These types of cells come in various sizes and thicknesses, all with similar appearances. They supply 1.5 V, and offer excellent energy storage for the weight. They also have a nearly flat discharge curve, like the one shown in the graph of Fig. 7-3. Zinc-carbon and alkaline cells and batteries, in contrast, have current output that declines more steadily with time, as shown in Fig. 7-5. This is known as a *declining discharge curve.*

Silver-oxide cells can be stacked to make batteries. Several of these miniature cells, one on top of the other, can provide 6, 9, or even 12 V for a transistor radio or other light-duty electronic device. The resulting battery is about the size of an AAA cylindrical cell.

Mercury Cells and Batteries

A *mercury cell*, also called a *mercuric-oxide cell*, has properties similar to those of silver-oxide cells. They are manufactured in the same general form. The main difference, often not of significance, is a somewhat lower voltage per cell: 1.35 V. If six of these cells are stacked to make a battery, the re-



7-5 A declining discharge curve.

sulting voltage will be about 8.1 V rather than 9 V. One additional cell can be added to the stack, yielding about 9.45 V.

There has been a decline in the popularity of mercury cells and batteries in recent years, because of the fact that mercury is toxic to humans and animals, even in trace amounts. When mercury cells and batteries are dead, they must be discarded. Eventually the mercury or mercuric oxide leaks into the soil and groundwater. Mercury pollution has become a significant concern throughout the world.

Lithium Cells and Batteries

Lithium cells gained popularity in the early 1980s. There are several variations in the chemical makeup of these cells; they all contain lithium, a light, highly reactive metal. Lithium cells can be made to supply 1.5 to 3.5 V, depending on the particular chemistry used. These cells, like silver-oxide and mercury cells, can be stacked to make batteries.

The first application of lithium batteries was in memory backup for electronic microcomputers. Lithium cells and batteries have superior shelf life, and they can last for years in very-low-current applications such as memory backup or the powering of a digital liquid crystal display (LCD) watch or clock. These cells also provide high energy capacity per unit volume or mass.

Lead-Acid Batteries

You've seen the basic configuration for a lead-acid cell. This has a solution of sulfuric acid, along with a lead electrode (negative) and a lead-dioxide electrode (positive). These cells are rechargeable.

Automotive batteries are made from sets of lead-acid cells having a free-flowing liquid acid. You cannot tip such a battery on its side, or turn it upside-down, without running the risk of having some of the acid electrolyte spill out. Lead-acid batteries are also available in a construction that uses a semisolid electrolyte. These batteries are sometimes used in consumer electronic devices that require a moderate amount of current. The most common example is an uninterruptible power supply (UPS) that can keep a desktop personal computer running for a few minutes if the utility power fails.

A large lead-acid battery, such as the kind in your car or truck, can store several tens of ampere-hours. The smaller ones, like those in a UPS, have less capacity but more versatility. Their main attributes are that they can be charged and recharged many times, and they are not particularly expensive.

Nickel-Based Cells and Batteries

Nickel-based cells include the *nickel-cadmium* (NICAD or NiCd) type and the *nickel-metal-hydride* (NiMH) type. *Nickel-based batteries* are available in packs of cells. These packs can be plugged into equipment, and sometimes form part of the case for a device such as a portable radio transceiver. All nickel-based cells are rechargeable, and can be put through hundreds or even thousands of *charge/discharge cycles* if they are properly cared for.

Configurations and Applications

Nickel-based cells are found in various sizes and shapes. *Cylindrical cells* look like ordinary dry cells. *Button cells* are those little things you find in cameras, watches, memory backup applications, and other places where miniaturization is important. *Flooded cells* are used in heavy-duty applications, and can have storage capacity in excess of 1000 Ah. *Spacecraft cells* are made in packages that can withstand the rigors of a deep-space environment.

Most orbiting satellites are in darkness half the time and in sunlight half the time. *Solar panels* can be used while the satellite is in sunlight, but during the times that the earth eclipses the sun, batteries are needed to power the electronic equipment on board the satellite. The solar panels can charge a nickel-based battery, in addition to powering the satellite, for the daylight half of each orbit. The nickel-based battery can provide the power during the dark half of each orbit.

Cautions

Never discharge nickel-based cells all the way until they totally die. This can cause the polarity of a cell, or of one or more cells in a battery, to reverse. Once this happens, the cell or battery is ruined.

A phenomenon peculiar to nickel-based cells and batteries is known as *memory* or *memory drain.* If a nickel-based unit is used over and over, and is discharged to the same extent every time, it might begin to die at that point in its discharge cycle. Memory problems can usually be solved. Use the cell or battery almost all the way up, and then fully recharge it. Repeat the process several times.

Nickel-based cells and batteries work best if used with charging units that take several hours to fully replenish the charge. So-called high-rate or quick chargers are available, but these can sometimes force too much current through a cell or battery. It's best if the charger is made especially for the cell or battery type being charged. An electronics dealer, such as the manager at a RadioShack store, should be able to tell you which chargers are best for which cells and batteries.

In recent years, concern has grown about the toxic environmental effects of discarded heavy metals, including cadmium. For this reason, NiMH cells and batteries have replaced NICAD types in many applications. In most practical scenarios, a NICAD battery can be directly replaced with a NiMH battery of the same voltage and current-delivering capacity, and the powered-up device will work satisfactorily.

Some vendors and dealers will call a nickel-based cell or battery a NICAD, even when it is actually a NiMH cell or battery.

Photovoltaic Cells and Batteries

The *photovoltaic* (PV) *cell* is different from any electrochemical cell. It's also known as a *solar cell.* This device converts visible light, infrared (IR), and/or ultraviolet (UV) directly into electric current.

Solar Panels

Several, or many, photovoltaic cells can be combined in series-parallel to make a *solar panel.* An example is shown in Fig. 7-6. Although this shows a 3×3 series-parallel array, the matrix does not have to be symmetrical. And it's often very large. It might consist of, say, 50 parallel sets of 20 series-connected cells. The series scheme boosts the voltage to the desired level, and the parallel scheme increases the current-delivering ability of the panel. It's not unusual to see hundreds of solar cells combined in this way to make a large panel.

Construction and Performance

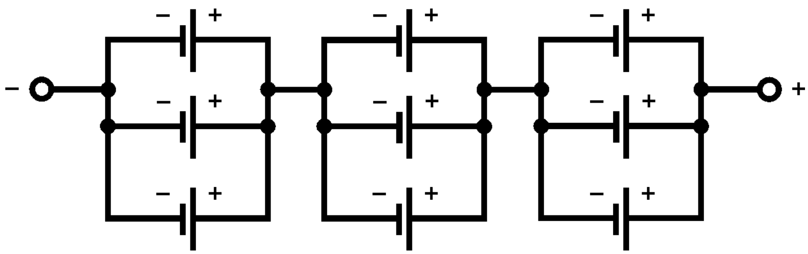
The construction of a photovoltaic cell is shown in Fig. 7-7. The device is a flat semiconductor *P-N junction*, and the assembly is made transparent so that light can fall directly on the *P-type* silicon. The metal ribbing, forming the positive electrode, is interconnected by means of tiny wires. The negative electrode is a metal backing or *substrate*, placed in contact with the *N-type* silicon.

Most solar cells provide about 0.5 V. If there is very low current demand, dim light will result in the full-output voltage from a solar cell. As the current demand increases, brighter light is needed to produce the full-output voltage. There is a maximum limit to the current that can be provided from a solar cell, no matter how bright the light. This limit is increased by connecting solar cells in parallel.

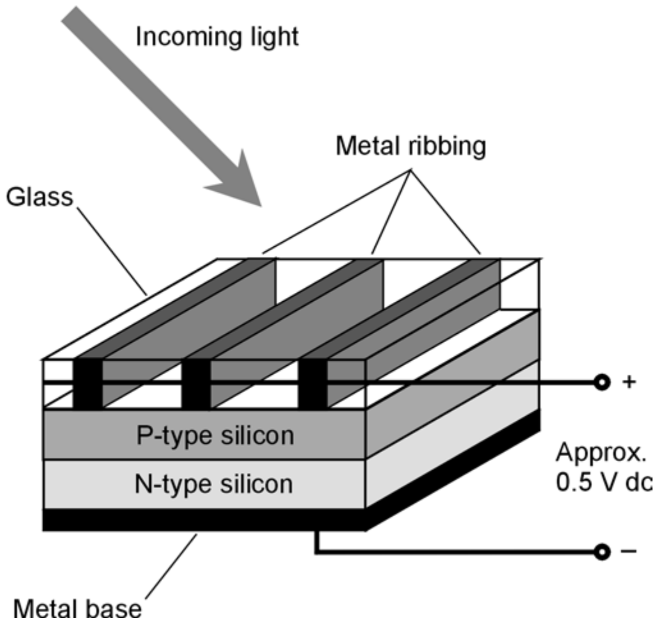
Practical Applications

Solar cells have become cheaper and more efficient in recent years, as researchers have looked to them as an alternative energy source. Solar panels are used in satellites. They can be used in conjunction with rechargeable batteries, such as the lead-acid or nickel-cadmium types, to provide power independent of the commercial utilities.

A completely independent solar/battery power system is called a *stand-alone* system. It uses large solar panels, large-capacity lead-acid batteries, power converters to convert the dc into ac, and a sophisticated charging circuit. These systems are best suited to environments where there is sunshine a high percentage of the time.



7-6 Connection of cells in series-parallel.



7-7 Construction of a silicon photovoltaic cell.

Solar cells, either alone or supplemented with rechargeable batteries, can be connected into a home electric system in an *interactive* arrangement with the electric utilities. When the solar power system can't provide for the needs of the household all by itself, the utility company can take up the slack. Conversely, when the solar power system supplies more than enough for the needs of the home, the utility company can buy the excess.

Fuel Cells

In the late 1900s, a new type of electrochemical power device emerged that is believed by some scientists and engineers to hold promise as an alternative energy source: the *fuel cell.*

Hydrogen Fuel

The most talked-about fuel cell during the early years of research and development became known as the *hydrogen fuel cell.* As its name implies, it derives electricity from hydrogen. The hydrogen combines with oxygen (that is, it *oxidizes*) to form energy and water. There is no pollution, and there are no toxic by-products. When a hydrogen fuel cell "runs out of juice," all that is needed is a new supply of hydrogen, because its oxygen is derived from the atmosphere.

Instead of combusting, the hydrogen in a fuel cell oxidizes in a more controlled fashion, and at a much lower temperature. There are several schemes for making this happen. The *proton exchange membrane* (PEM) *fuel cell* is one of the most widely used. A PEM hydrogen fuel cell generates approximately 0.7 V of dc. In order to obtain higher voltages, individual cells are connected in series. A series-connected set of fuel cells is technically a battery, but the term used more often is *stack.*

Fuel-cell stacks are available in various sizes. A stack about the size and weight of an airline suitcase filled with books can power a subcompact electric car. Smaller cells, called *micro fuel cells*, can provide dc to run devices that have historically operated from conventional cells and batteries. These include portable radios, lanterns, and notebook computers.

Other Fuels

Hydrogen is not the only chemical that can be used to make a fuel cell. Almost anything that will combine with oxygen to form energy has been considered.

Methanol, a form of alcohol, has the advantage of being easier to transport and store than hydrogen, because it exists as a liquid at room temperature. *Propane* is another chemical that has been used for powering fuel cells. This is the substance that is stored in liquid form in tanks for barbecue grills and some rural home heating systems. *Methane*, also known as natural gas, has been used as well.

Some scientists and engineers object to the use of these fuels because they, especially propane and methane, closely resemble fuels that are already commonplace, and on which society has developed the sort of dependence that purists would like to get away from. In addition, they are derived from so-called *fossil fuel* sources, the supplies of which, however great they might be today, are nevertheless finite.

A Promising Technology

As of this writing (2006), fuel cells have not yet replaced conventional electrochemical cells and batteries. Cost is the main reason. Hydrogen is the most abundant and simplest chemical element in the universe, and it does not produce any toxic by-products. This would at first seem to make it the ideal choice for use in fuel cells. But storage and transport of hydrogen has proven to be difficult and expensive. This is especially true for fuel cells and stacks intended for systems that aren't fixed to permanent pipelines.

An interesting scenario, suggested by one of my physics teachers all the way back in the 1970s, is the piping of hydrogen gas through the lines designed to carry methane. Some modification of existing lines would be required in order to safely handle hydrogen, which escapes through small cracks and openings more easily than methane. But hydrogen, if obtained at reasonable cost and in abundance, could be used to power large fuel-cell stacks in common households and businesses. The dc from such a stack could be converted to utility ac by power inverters similar to those used with PV energy systems. The entire home power system would be about the size of a gas furnace.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. The chemical energy in a battery or cell
   (a) is a form of kinetic energy.
   (b) cannot be replenished once it is gone.
   (c) changes to electrical energy when the cell is used.
   (d) is caused by electric current.

2. A cell that cannot be recharged is known as
   (a) a dry cell.
   (b) a wet cell.
   (c) a primary cell.
   (d) secondary cell.
3. A Weston cell is generally used
   (a) as a current reference source.
   (b) as a voltage reference source.
   (c) as a power reference source.
   (d) as a fuel cell.
4. The voltage produced by a battery of multiple cells connected in series is
   (a) less than the voltage produced by a cell of the same composition.
   (b) the same as the voltage produced by a cell of the same composition.
   (c) more than the voltage produced by a cell of the same composition.
   (d) always a whole-number multiple of 1.018 V.
5. A direct short-circuit of a large battery can cause
   (a) an increase in its voltage.
   (b) no harm other than a rapid discharge of its energy.
   (c) the current to drop to zero.
   (d) a physical rupture or explosion.
6. Suppose a cell of 1.5 V delivers 100 mA for 7 hours and 20 minutes, and then it is replaced. How much energy is supplied during this time?
   (a) 0.49 Wh
   (b) 1.1 Wh
   (c) 7.33 Wh
   (d) 733 mWh
7. Suppose a 12-V automotive battery is rated at 36 Ah. If a 100-W, 12-V bulb is connected across this battery, approximately how long will the bulb stay aglow, assuming the battery has been fully charged?
   (a) 4 hours and 20 minutes
   (b) 432 hours
   (c) 3.6 hours
   (d) 21.6 minutes
8. Alkaline cells
   (a) are cheaper than zinc-carbon cells.
   (b) generally work better in radios than zinc-carbon cells.
   (c) have higher voltages than zinc-carbon cells.
   (d) have shorter shelf lives than zinc-carbon cells.

9. The energy in a cell or battery depends mainly on
   (a) its physical size.
   (b) the current drawn from it.
   (c) its voltage.
   (d) all of the above.

10. In which of the following devices would a lantern battery most likely be found?
    (a) A heart pacemaker
    (b) An electronic calculator
    (c) An LCD wall clock
    (d) A two-way portable radio

11. In which of the following devices would a transistor battery be the best power choice?
    (a) A heart pacemaker
    (b) An electronic calculator
    (c) An LCD wall clock
    (d) A two-way portable radio

12. For which of the following applications would you choose a lithium battery?
    (a) A microcomputer memory backup
    (b) A two-way portable radio
    (c) A stand-alone solar-electric system
    (d) A rechargeable lantern

13. Where would you most likely find a lead-acid battery?
    (a) In a portable audio CD player
    (b) In an uninterruptible power supply
    (c) In an LCD wall clock
    (d) In a flashlight

14. A cell or battery that maintains a constant current-delivering capability almost until it dies is said to have
    (a) a large ampere-hour rating.
    (b) excellent energy capacity.
    (c) a flat discharge curve.
    (d) good energy storage capacity per unit volume.

15. Where might you find a nickel-based battery?
    (a) In a satellite
    (b) In a portable cassette player
    (c) In a handheld radio transceiver
    (d) More than one of the above

16. A disadvantage of mercury cells and batteries is the fact that
    (a) they don't last as long as other types.
    (b) they have a flat discharge curve.
    (c) mercury is destructive to the environment.
    (d) they need to be recharged often.
17. Which kind of battery should never be used until it dies?
    (a) Silver-oxide
    (b) Lead-acid
    (c) Nickel-based
    (d) Mercury
18. The useful current that is delivered by a solar panel can be increased by
    (a) connecting capacitors in parallel with the solar cells.
    (b) connecting resistors in series with the solar cells.
    (c) connecting two or more groups of solar cells in parallel.
    (d) connecting resistors in parallel with the solar cells.
19. An interactive solar power system
    (a) allows a homeowner to sell power to the electric company.
    (b) lets the batteries recharge at night.
    (c) powers lights, but not electronic devices.
    (d) is totally independent from the electric company.
20. An advantage of methanol over hydrogen for use in fuel cells is the fact that
    (a) methanol is the most abundant element in the universe.
    (b) methanol is not flammable.
    (c) methanol is a solid at room temperature.
    (d) methanol is easier to transport and store.

8
CHAPTER

Magnetism

ELECTRIC AND MAGNETIC PHENOMENA INTERACT. MAGNETISM WAS MENTIONED BRIEFLY NEAR THE end of Chap. 2. Here, we'll look at it more closely.

The Geomagnetic Field

The earth has a core made up largely of iron, heated to the extent that some of it is liquid. As the earth rotates, the iron flows in complex ways. It is thought that this flow is responsible for the magnetic field that surrounds the earth. Some other planets, notably Jupiter, have magnetic fields as well. Even the sun has one.

The Poles and Axis

The *geomagnetic field*, as it is called, has poles, just as a bar magnet does. The *geomagnetic poles* are near, but not at, the *geographic poles.* The *north geomagnetic pole* is located in far northern Canada. The *south geomagnetic pole* is near Antarctica. The *geomagnetic axis* is therefore tilted relative to the axis on which the earth rotates.

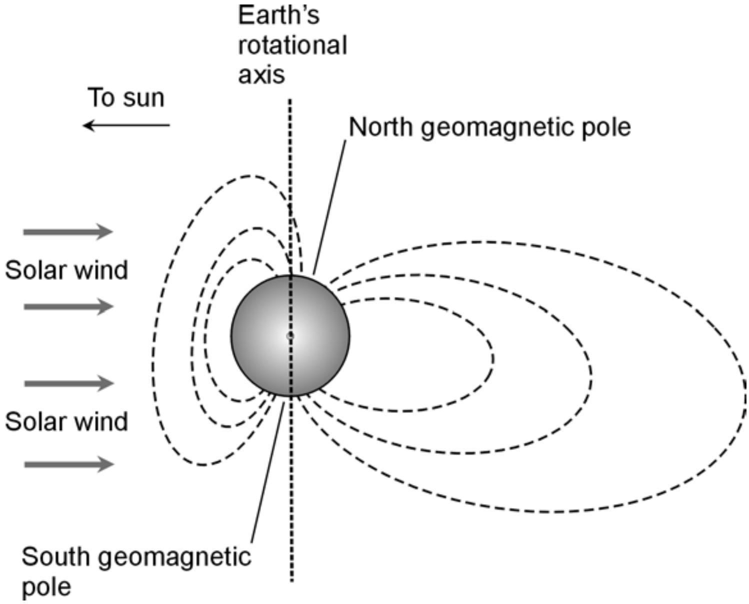
The Solar Wind

Charged subatomic particles from the sun, streaming outward through the solar system, distort the geomagnetic *lines of flux* (Fig. 8-1). This stream of particles is called the *solar wind.* That's a good name for it, because the fast-moving particles produce measurable forces on sensitive instruments in space. This force has actually been suggested as a possible means to drive space ships, equipped with *solar sails*, out of the solar system!

At and near the earth's surface, the geomagnetic field is not affected very much by the solar wind, so the geomagnetic field is nearly symmetrical. As the distance from the earth increases, the distortion of the field also increases, particularly on the side of the earth away from the sun.

The Magnetic Compass

The presence of the geomagnetic field was first noticed in ancient times. Some rocks, called *lodestones*, when hung by strings, would always orient themselves a certain way. This was correctly at-



8-1 Geomagnetic flux lines (dashed curves) are distorted by the solar wind, so the geomagnetic field is not symmetrical with respect to the earth.

tributed to the presence of a "force" in the air. This effect was put to use by early seafarers and land explorers. Today, a *magnetic compass* can still be a valuable navigation aid, used by mariners, backpackers, and others who travel far from familiar landmarks.

The geomagnetic field interacts with the magnetic field around a compass needle, and a force is thus exerted on the needle. This force works not only in a horizontal plane (parallel to the earth's surface), but vertically at most latitudes. The vertical component is zero only at the *geomagnetic equator*, a line running around the globe equidistant from both geomagnetic poles.

As the geomagnetic latitude increases, toward either the north or the south geomagnetic pole, the magnetic force pulls up and down on the compass needle more and more. One end of the needle seems to insist on touching the compass face, while the other end tilts up toward the glass. The needle tries to align itself parallel to the geomagnetic *lines of flux.* The vertical angle, in degrees, at which the geomagnetic lines of flux intersect the earth's surface at any given location is called the *geomagnetic inclination.*

Because geomagnetic north is not the same as geographic north in most places on the earth's surface, there is an angular difference between the two. This horizontal angle, in degrees, is called *geomagnetic declination.* It, like inclination, varies with location.

Causes and Effects

Magnets are attracted to some, but not all, metals. Iron, nickel, and alloys containing either or both of these elements are known as *ferromagnetic* materials. They "stick" to magnets. They can also be made into *permanent magnets.* When a magnet is brought near a piece of ferromagnetic material, the atoms in the material become lined up, so that the material is temporarily magnetized. This produces a *magnetic force* between the atoms of the ferromagnetic substance and those in the magnet.

Attraction and Repulsion

If a magnet is brought near another magnet, the force can be repulsive or attractive, depending on the way the magnets are oriented. The force gets stronger as the magnets are brought near each other. Some magnets are so strong that no human being can pull them apart if they get stuck together, and no person can bring them all the way together against their mutual repulsive force. This is especially true of *electromagnets*, discussed later in this chapter.

The tremendous forces produced by electromagnets are of use in industry. A large electromagnet can be used to carry heavy pieces of scrap iron from place to place. Other electromagnets can provide sufficient repulsion to suspend one object above another. This phenomenon is called *magnetic levitation.* It is the basis for low-friction, high-speed commuter trains now in use in some metropolitan areas.

Charge in Motion

Whenever the atoms in a ferromagnetic material are aligned, a *magnetic field* exists. A magnetic field can also be caused by the motion of electric charge carriers, either in a wire or in free space.

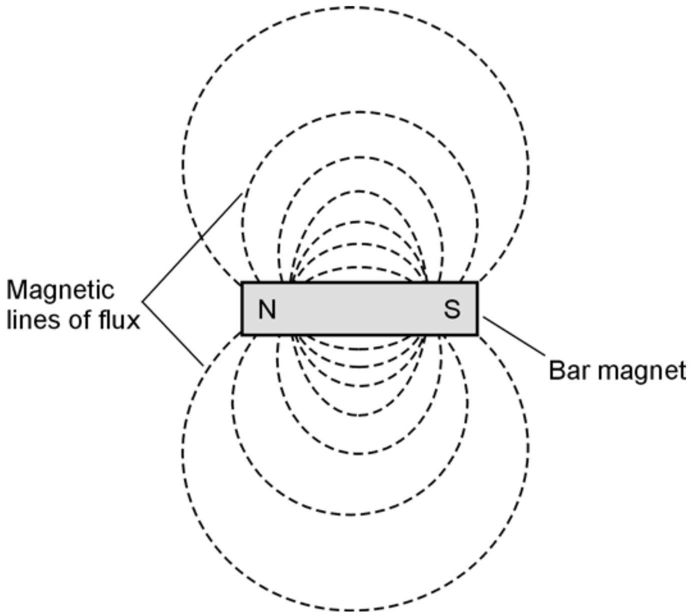
The magnetic field around a permanent magnet arises from the same cause as the field around a wire that carries an electric current. The responsible factor in either case is the motion of electrically charged particles. In a wire, electrons move along the conductor, being passed from atom to atom. In a permanent magnet, the movement of orbiting electrons occurs in such a manner that an effective electrical current is produced.

Magnetic fields are also generated by the motion of charged particles through space. The sun is constantly ejecting protons and helium nuclei. These particles carry a positive electric charge. Because of this, and the fact that they are in motion, they are surrounded by tiny magnetic fields. When the particles approach the earth and their magnetic fields interact with the geomagnetic field, the particles are accelerated toward the geomagnetic poles.

When there is a *solar flare*, the sun ejects far more charged particles than normal. When these approach the geomagnetic poles, the result is considerable disruption of the geomagnetic field. This type of event is called a *geomagnetic storm.* It causes changes in the earth's ionosphere, affecting long-distance radio communications at certain frequencies. If the fluctuations are intense enough, even wire communications and electric power transmission can be interfered with. Aurora (northern or southern lights) are frequently observed at night during these events.

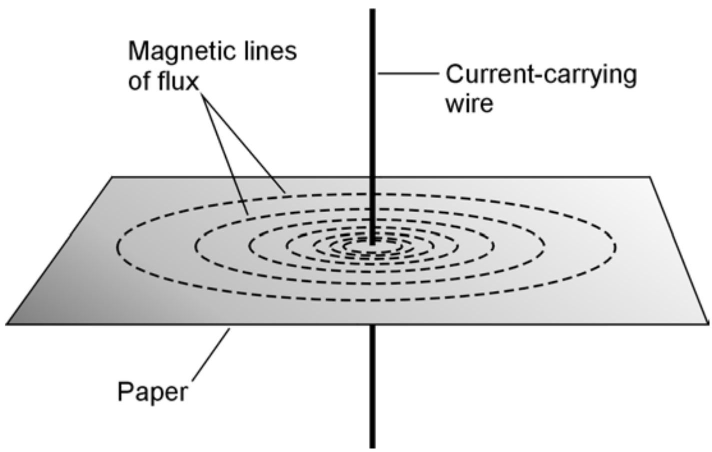
Flux Lines

Have you seen the well-known experiment in which iron filings are placed on a horizontal sheet of paper, and then a magnet is placed underneath the paper? The filings arrange themselves in a pattern that shows, roughly, the shape of the magnetic field in the vicinity of the magnet. A bar magnet has a field with a characteristic form (Fig. 8-2). Another popular experiment involves passing a current-carrying wire through a horizontal sheet of paper at a right angle, as shown in Fig. 8-3. The iron filings become grouped along circles centered at the point where the wire passes through the paper.



8-2 The pattern of magnetic flux lines (dashed curves) around a bar magnet (rectangle). The N and S represent north and south magnetic poles, respectively.

The intensity of a magnetic field is determined according to the number of flux lines passing through a certain cross section, such as a square centimeter or a square meter. The lines don't exist as real objects, but it is intuitively appealing to imagine them that way. The iron filings on the paper really do bunch themselves into lines (curves, actually) when there is a magnetic field of sufficient strength to make them move. Sometimes lines of flux are called *lines of force.* But technically, this is a misnomer.



8-3 The pattern of magnetic flux lines (dashed curves) around a straight, current-carrying wire can be seen when the wire passes through a horizontal sheet of paper sprinkled with iron filings.

Poles

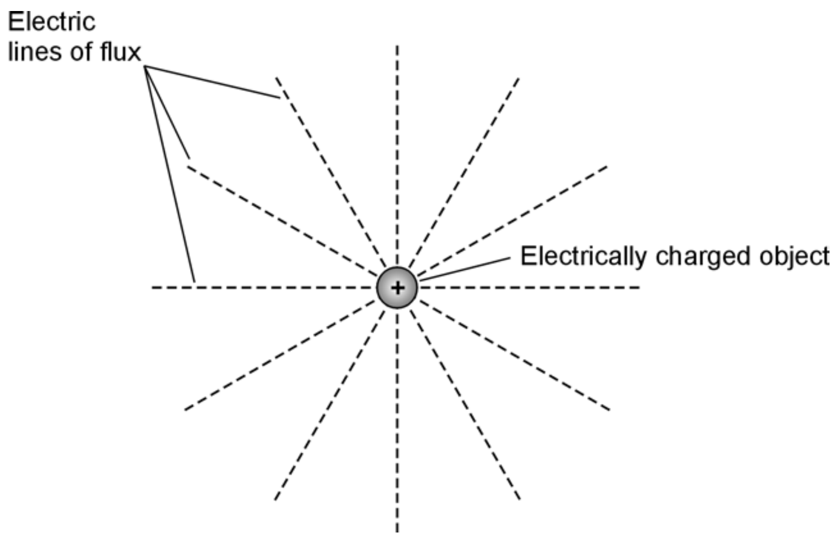
A magnetic field has a specific direction, as well as a specific intensity, at any given point in space near a current-carrying wire or a permanent magnet. The flux lines run parallel with the direction of the field. A magnetic field is considered to begin at the north magnetic pole, and to terminate at the south magnetic pole. In the case of a permanent magnet, it is obvious where the magnetic poles are. In the case of a current-carrying wire, the magnetic field goes in endless circles around the wire.

A charged electric particle, such as a proton or electron, hovering all by itself in space, constitutes an *electric monopole.* The *electric lines of flux* around an isolated, charged particle in free space are straight, and they "run off to infinity" (Fig. 8-4). A positive electric charge does not have to be mated with a negative electric charge.

A magnetic field is different. All magnetic flux lines, at least in ordinary real-world situations, are *closed loops.* With permanent magnets, there is a starting point (the north pole) and an ending point (the south pole). Around a straight, current-carrying wire, the loops are closed circles, even though the starting and ending points are not obvious. A pair of magnetic poles is called a *magnetic dipole.*

At first you might think that the magnetic field around a current-carrying wire is caused by a monopole, or that there aren't any poles at all, because the concentric circles don't actually converge anywhere. But you can envision a half plane, with the edge along the line of the wire, as a magnetic dipole. Then the lines of flux go around once in a 360° circle from the "north face" of the half plane to the "south face."

The greatest flux density, or field strength, around a bar magnet is near the poles, where the lines converge. Around a current-carrying wire, the greatest field strength is near the wire.



8-4 Electric flux lines (dashed lines) around an electrically charged object. This example shows a positive charge. The pattern of flux lines for a negative charge is identical.

Magnetic Field Strength

The overall magnitude of a magnetic field is measured in units called *webers* (Wb). A smaller unit, the *maxwell* (Mx), is sometimes used if a magnetic field is weak. One weber is equivalent to 100,000,000 (10^8) maxwells. Conversely, 1 Mx = 0.00000001 Wb = 10^{-8} Wb.

The Tesla and the Gauss

If you have access to a permanent magnet or electromagnet, you might see its strength expressed in terms of webers or maxwells. But usually you'll hear units called *teslas* (T) or *gauss* (G). These units are expressions of the concentration, or intensity, of the magnetic field within a certain cross section. The *flux density*, or number of lines per square meter or per square centimeter, is a more useful expression for magnetic effects than the overall quantity of magnetism. A flux density of 1 tesla (1 T) is equal to 1 weber per square meter (1 Wb/m^2). A flux density of 1 gauss (1 G) is equal to 1 maxwell per square centimeter (1 Mx/cm^2). It turns out that the gauss is equal to 0.0001 tesla (10^{-4} T). Conversely, the tesla is equivalent to 10,000 gauss (10^4 G).

The Ampere-Turn and the Gilbert

With electromagnets, another unit is employed: the *ampere-turn* (At). This is technically a unit of *magnetomotive force*, which is the magnetic counterpart of electromotive force. A wire, bent into a circle and carrying 1 A of current, produces 1 At of magnetomotive force. If the wire is bent into a loop having 50 turns, and the current stays the same, the resulting magnetomotive force is 50 At. If the current is then reduced to 1/50 A or 20 mA, the magnetomotive force will go back down to 1 At.

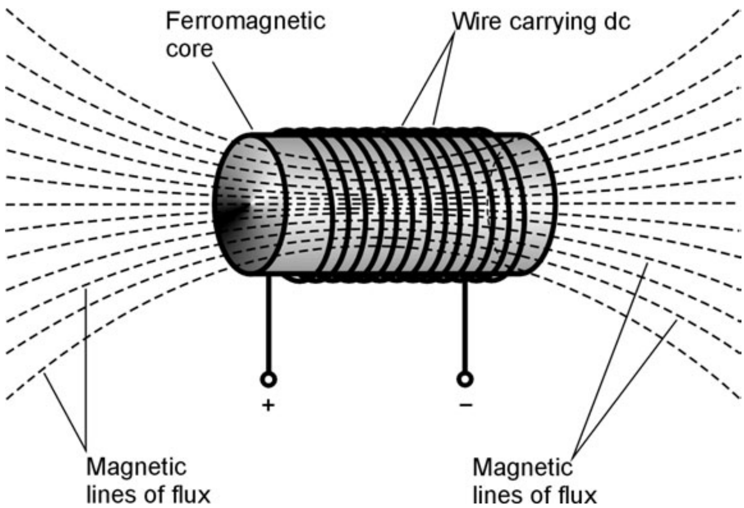
The *gilbert* (Gb) is also used to express magnetomotive force, but it is less common than the ampere-turn. One gilbert (1 Gb) is equal to 0.796 At. Conversely, 1 At = 1.26 Gb.

Electromagnets

Any electric current, or movement of charge carriers, produces a magnetic field. This field can become intense in a tightly coiled wire that has many turns and carries a large current. When a ferromagnetic core is placed inside the coil, the magnetic lines of flux are concentrated in the core, and the field strength in and near the core can become tremendous. This is the principle of an *electromagnet* (Fig. 8-5). Electromagnets are almost always cylindrical in shape. Sometimes the cylinder is long and thin; in other cases it is short and fat. But whatever the ratio of diameter to length for the core, the principle is the same: the magnetic field produced by the current results in magnetization of the core.

Direct-Current Types

You can build a dc electromagnet by taking a large bolt, such as a stove bolt, and wrapping a few dozen or a few hundred turns of wire around it. These items are available in any good hardware store. Be sure the bolt is made of ferromagnetic material. (If a permanent magnet sticks to the bolt, the bolt is ferromagnetic.) Ideally, the bolt should be at least 1 cm (approximately ⅜ in) in diameter and several inches long. You must use insulated wire, preferably made of solid, soft copper. "Bell wire" works well. Be sure all the wire turns go in the same direction. A large 6-V lantern battery can provide plenty of current to work the electromagnet. Never leave the coil connected to the battery for more than a few seconds at a time. And never use a car battery for this experiment! The acid can boil out of this type of battery, because the electromagnet places a heavy load on it.

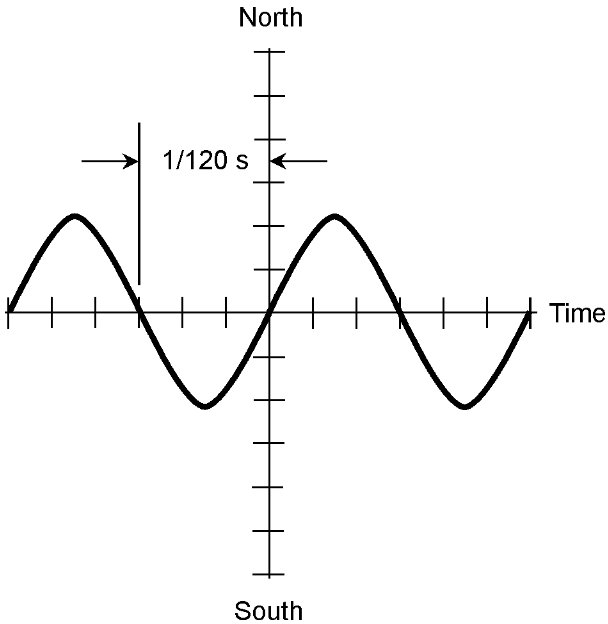


8-5 In an electromagnet, the magnetic flux is concentrated in a ferromagnetic rod surrounded by a current-carrying coil.

Direct-current electromagnets have defined north and south poles, just like permanent magnets. The main difference is that an electromagnet can get much stronger than any permanent magnet. You will see evidence of this if you do the preceding experiment with a large enough bolt and enough turns of wire.

Alternating-Current Types

Do you get the idea that an electromagnet can be made far stronger if, rather than using a lantern battery for the current source, you plug the wires into a wall outlet? In theory, this is true. In prac-



8-6 Polarity change in an ac electromagnet. The polarity changes every 1/120 second for 60-Hz utility current.

tice, you'll blow the fuse or circuit breaker. *Do not try this!* The electrical circuits in some buildings are not adequately protected and it can create a fire hazard. Also, you can get a lethal shock from the utility mains.

Some electromagnets use ac, and these magnets will stick to ferromagnetic objects. But the polarity of the magnetic field reverses every time the direction of the current reverses. With conventional household ac in the United States, there are 120 fluctuations, or 60 complete north-to-south-to-north polarity changes (Fig. 8-6), per second. If a permanent magnet, or a dc electromagnet, is brought near either "pole" of an ac electromagnet, there is no net force because the poles are alike half the time and opposite half the time, producing an equal amount of attractive and repulsive force. But if a piece of iron or steel is brought near a strong ac electromagnet, watch out! The attractive force will be powerful.

Magnetic Properties of Materials

There are four important properties that materials can have with respect to magnetic flux. These properties are *ferromagnetism*, *diamagnetism*, *permeability*, and *retentivity.*

Ferromagnetism

Some substances cause magnetic lines of flux to bunch closer together than they would in the medium of air or a vacuum. This property is called *ferromagnetism*, and materials that exhibit it are called *ferromagnetic.* You've already learned something about this!

Diamagnetism

Another property is known as *diamagnetism*, and materials that exhibit it are called *diamagnetic.* This type of substance decreases the magnetic flux density by causing the magnetic flux lines to diverge. Wax, dry wood, bismuth, and silver are examples. No diamagnetic material reduces the strength of a magnetic field by anywhere near the factor that ferromagnetic substances can increase it. Diamagnetic materials are generally used to keep magnetic objects apart, while minimizing the interaction between them. In recent years, they have also found some application in *magnetic levitation* devices.

Permeability

Permeability is a quantitative indicator of the extent to which a ferromagnetic material concentrates magnetic lines of flux. It is measured on a scale relative to a vacuum, or free space. Free space is assigned permeability 1. If you have a coil of wire with an air core, and a current is forced through the wire, then the flux in the coil core is at a certain density, just about the same as it would be in a vacuum. Therefore, the permeability of pure air is about equal to 1. If you place an iron core in the coil, the flux density increases by a large factor. The permeability of iron can range from 60 (impure) to as much as 8000 (highly refined).

If you use certain ferromagnetic alloys as the core material in electromagnets, you can increase the flux density, and therefore the local strength of the field, by as much as a million times. Such substances thus have permeability as great as 1,000,000 (10^6).

Table 8-1 gives permeability values for some common materials.

Retentivity

When a substance, such as iron, is subjected to a magnetic field as intense as it can handle, say by enclosing it in a wire coil carrying a massive current, there will be some *residual magnetism* left

Table 8-1. Permeability values for some common materials.

Substance	Permeability (approx.)
Air, dry, at sea level	1
Alloys, ferromagnetic	3000–1,000,000
Aluminum	Slightly more than 1
Bismuth	Slightly less than 1
Cobalt	60–70
Iron, powdered and pressed	100–3000
Iron, solid, refined	3000–8000
Iron, solid, unrefined	60–100
Nickel	50–60
Silver	Slightly less than 1
Steel	300–600
Vacuum	1
Wax	Slightly less than 1
Wood, dry	Slightly less than 1

when the current stops flowing in the coil. *Retentivity*, also sometimes called *remanence*, is a measure of how well the substance "memorizes" the magnetism and thereby becomes a permanent magnet.

Retentivity is expressed as a percentage, and is symbolized B_r. If the flux density in the material is x tesla or gauss when it is subjected to the greatest possible magnetomotive force, and then goes down to y tesla or gauss when the current is removed, the retentivity is equal to $100(y/x)\%$.

Suppose that a metal rod can be magnetized to 135 G when it is enclosed by a coil carrying an electric current. Imagine that this is the maximum possible flux density that the rod can be forced to have. (For any substance, there is always such a maximum.) Now suppose that the current is shut off, and 19 G remain in the rod. Then the retentivity, B_r, is calculated as follows:

$$B_r = 100(19/135)\% = (100 \times 0.14)\% = 14\%$$

Some ferromagnetic substances have high retentivity. These materials are excellent for making permanent magnets. Other substances have low retentivity. They work well as electromagnets, but not as permanent magnets.

If a ferromagnetic substance has poor retentivity, it is especially well-suited for use as the core material for an ac electromagnet, because the polarity of the magnetic flux can reverse within the material at a rapid rate. Materials with high retentivity do not work well for ac electromagnets, because they resist the polarity reversal that takes place with ac.

Practical Magnetism

Magnetism has numerous applications in common consumer devices and systems. Here are some of the more common ways in which magnetic phenomena can be put to use.

Permanent Magnets

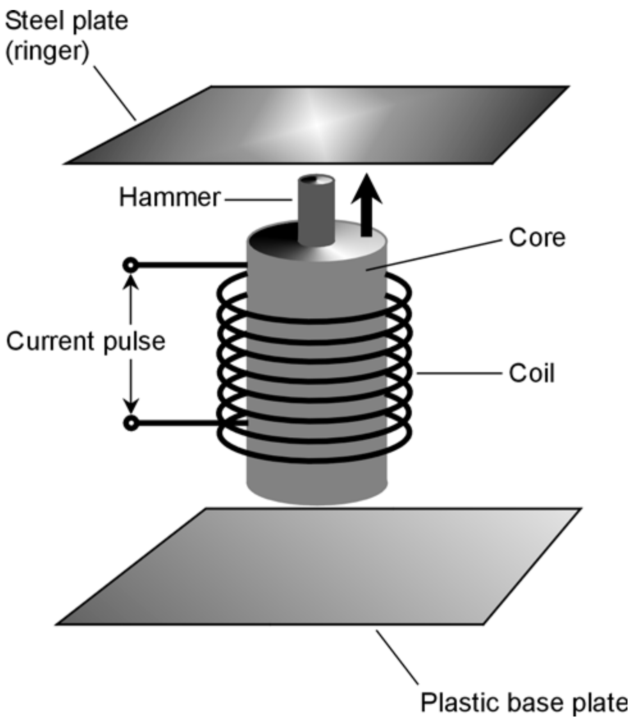
Permanent magnets are manufactured by using a high-retentivity ferromagnetic material as the core of an electromagnet for an extended period of time. The coil of the electromagnet carries a large direct current, causing intense magnetic flux of constant polarity within the material. (Don't try to do this at home. The high current can heat the coil and overload a battery or power supply, which produces a fire hazard and/or the risk of battery explosion.)

If you want to magnetize a screwdriver a little bit so that it will hold onto screws, just stroke the shaft of the screwdriver with the end of a bar magnet several dozen times. Once you have magnetized a tool in this way, however, it is nearly impossible to demagnetize it.

A Ringer Device

Figure 8-7 is a simplified diagram of a *bell ringer*, also called a *chime.* The main functional component is called a *solenoid*, and it is an electromagnet. The core has a hole going along its axis. The coil has several layers, but the wire is always wound in the same direction, so that the electromagnet is powerful. A movable steel rod runs through the hole in the electromagnet core.

When there is no current flowing in the coil, the steel rod is held down by the force of gravity. When a pulse of current passes through the coil, the rod is pulled forcibly upward so that it strikes the ringer plate. This plate is like one of the plates in a xylophone. The current pulse is short, so the steel rod falls back down again to its resting position, allowing the plate to reverberate.



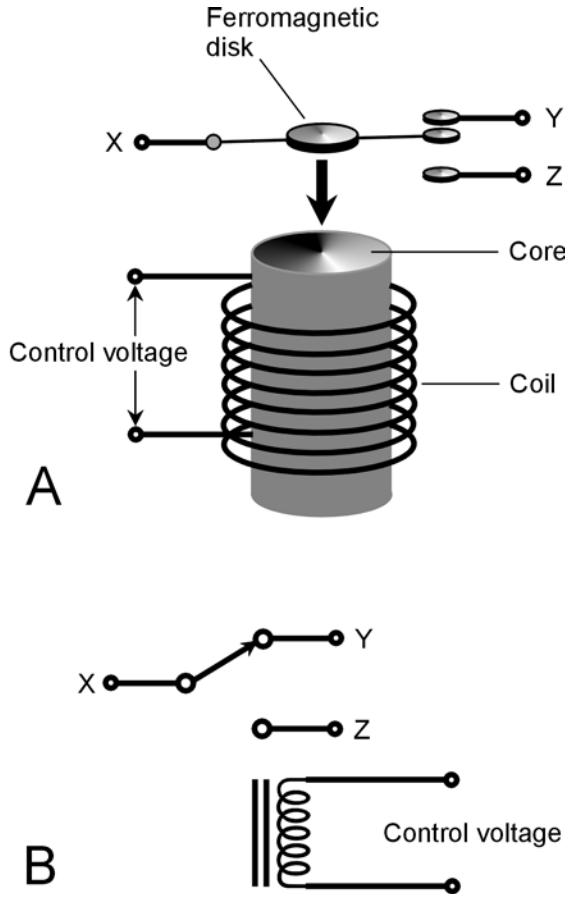
8-7 A solenoid-coil bell ringer.

The Relay

A *relay* makes use of a solenoid to allow remote-control switching of high-current circuits. A diagram of a relay is shown in Fig. 8-8. The movable lever, called the *armature*, is held to one side by a spring when there is no current flowing through the electromagnet. Under these conditions, terminal X is connected to Y, but not to Z. When a sufficient current is applied, the armature is pulled over to the other side. This disconnects terminal X from terminal Y, and connects X to Z.

There are numerous types of relays. Some are meant for use with dc, and others are for ac; a few will work with either dc or ac. A *normally closed relay* completes the circuit when there is no current flowing in its electromagnet coil, and breaks the circuit when current flows through the coil. A *normally open relay* is just the opposite, completing the circuit when current flows through the electromagnet coil, and opening the circuit when current ceases to flow through the coil. *Normal*, in this context, refers to the condition of no current applied to the electromagnet.

The relay shown in Fig. 8-8 can be used as either a normally open or normally closed relay, depending on which contacts are selected. It can also be used to switch a line between two different circuits.



8-8 At A, pictorial diagram of a simple relay. At B, the schematic symbol for the same relay.

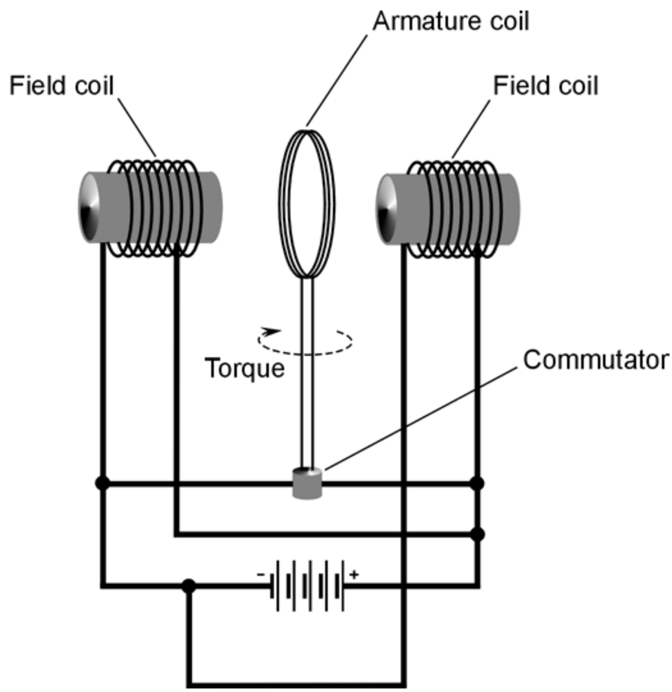
Some relays have several sets of contacts. Some relays are meant to remain in one state (either with current or without) for a long time, while others are meant to switch several times per second. The fastest relays can operate several dozen times per second. In recent years, relays have been largely supplanted by switching transistors and diodes, except in applications where extremely high current or high voltage is involved.

The DC Motor

Magnetic forces can be harnessed to do work. One common device that converts direct-current energy into rotating mechanical energy is a *dc motor.* In a dc motor, the source of electricity is connected to a set of coils, producing magnetic fields. The attraction of opposite poles, and the repulsion of like poles, is switched in such a way that a constant *torque*, or rotational force, results. As the current in the coils increases, the torque that the motor can provide also increases.

Figure 8-9 is a simplified, cutaway drawing of a dc motor. One set of coils, called the *armature coil*, rotates along with the motor shaft. The other set of coils, called the *field coil*, is stationary. The current direction is periodically reversed during each rotation by means of the *commutator.* This keeps the rotational force going in the same angular direction, so the motor continues to rotate rather than oscillating back and forth. The shaft is carried along by its own inertia, so that it doesn't come to a stop during those instants when the current is being switched in polarity.

Some dc motors can also be used to generate dc. These motors contain permanent magnets in place of one of the sets of coils. When the shaft is rotated, a pulsating dc flows in the coil.



8-9 A functional diagram of a dc motor.

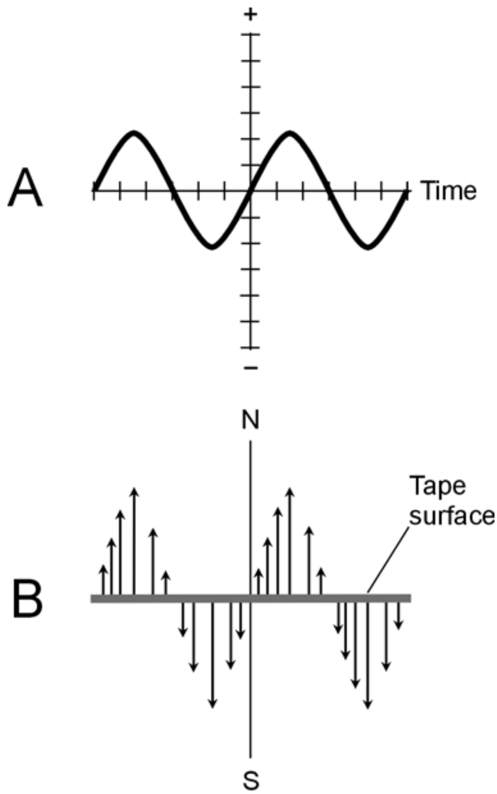
Magnetic Tape

Magnetic tape, also called *recording tape*, consists of millions of ferromagnetic particles attached to a flexible, thin plastic strip. In the *tape recorder*, a fluctuating magnetic field, produced by the *recording head*, polarizes these particles. As the field changes in strength next to the recording head, the tape passes by at a constant speed. This produces regions in which the ferromagnetic particles are polarized in either direction (Fig. 8-10).

When the tape is run at the same speed through the recorder in the playback mode, the magnetic fields around the individual particles cause a fluctuating field that is detected by the *pickup head.* This field has the same pattern of variations as the original field from the recording head.

Magnetic tape is available in various widths and thicknesses. Thicker tapes result in cassettes that don't play as long, but the tape is more resistant to stretching. The speed of the tape determines the fidelity of the recording. Higher speeds are preferred for music and video, and lower speeds for voice and data.

The impulses on a magnetic tape can be distorted or erased by external magnetic fields. Therefore, tapes should be protected from such fields. Keep the tape away from magnets. Extreme heat can also result in loss of data, and can cause permanent physical damage to the tape.



8-10 On recording tape, particles are magnetized in a pattern that follows the input waveform. Graph A shows an example of an audio input waveform. Graph B shows relative polarity and intensity of magnetization for selected particles on the tape surface.

Magnetic Disk

Since the advent of the personal computer, ever-more compact data-storage systems have evolved. One of the most versatile is the *magnetic disk.*

Hard disks, also called *hard drives*, store the most data, and are generally found inside of computer units. *Diskettes* are 8.9 cm (3.5 in) across, and can be inserted and removed from recording/playback machines called *diskette drives.* In recent years, magnetic diskettes have been largely supplanted by non-magnetic *compact disc recordable* (CD-R) and *compact disc rewritable* (CD-RW) media.

The principle of the magnetic disk, on the microscale, is the same as that of magnetic tape. The information is stored in *binary digital* form; that is, there are only two different ways that the particles are magnetized. This results in almost perfect, error-free storage. On a larger scale, the disk works differently than the tape because of the difference in geometry. On a tape, the information is spread out over a long span, and some bits of data are far away from others as measured along the medium itself. But on a disk, no two bits are ever farther apart than the diameter of the disk. This means that data can be stored to, and retrieved from, a disk much faster than is possible with tape.

The same precautions should be observed when handling and storing magnetic disks as are necessary with magnetic tape.

Bubble Memory

Bubble memory is a sophisticated method of storing data that gets rid of the need for moving parts such as are required in tape machines and disk drives. Data is stored as tiny magnetic fields, in a medium that is made from magnetic film and semiconductor materials.

Bubble memory makes use of all the advantages of magnetic data storage, as well as the favorable aspects of electronic data storage. Advantages of electronic memory include rapid storage and recovery, and high density (a lot of data can be put in a tiny volume of space). Advantages of magnetic memory include *nonvolatility* (it can be stored for a long time without needing a constant current source), high density, and comparatively low cost.

Bubble memory seems to go through phases. Just as it is declared obsolete, someone comes up with a new and improved way to make it work. Check the Internet to find out its current status; enter "bubble memory" or "magnetic bubble memory" into a search engine.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct. Answers are in the back of the book.

1. The geomagnetic field
   (a) makes the earth like a huge horseshoe magnet.
   (b) runs exactly through the geographic poles.
   (c) makes a compass work.
   (d) makes an electromagnet work.

2. Geomagnetic lines of flux
   (a) are horizontal at the geomagnetic equator.
   (b) are vertical at the geomagnetic equator.

(c) are never horizontal, no matter where you go.
(d) are perfectly symmetrical around the earth, even far out in space.

3. A material that can be permanently magnetized is generally said to be
(a) ultramagnetic.
(b) electromagnetic.
(c) diamagnetic.
(d) ferromagnetic.

4. The force between a magnet and a piece of ferromagnetic metal that has not been magnetized
(a) can be either repulsive or attractive.
(b) is never repulsive.
(c) gets smaller as the magnet gets closer to the metal.
(d) depends on the geomagnetic field.

5. The presence of a magnetic field can always be attributed to
(a) ferromagnetic materials.
(b) diamagnetic materials.
(c) motion of electric charge carriers.
(d) the north geomagnetic pole.

6. Lines of magnetic flux are said to originate
(a) in atoms of ferromagnetic materials.
(b) at a north magnetic pole.
(c) at points where the lines are straight.
(d) in electric charge carriers.

7. The magnetic flux around a straight, current-carrying wire
(a) gets stronger with increasing distance from the wire.
(b) is strongest near the wire.
(c) does not vary in strength with distance from the wire.
(d) consists of straight lines parallel to the wire.

8. The gauss is a unit of
(a) overall magnetic field strength.
(b) ampere-turns.
(c) magnetic flux density.
(d) magnetic power.

9. A unit of overall magnetic field quantity is the
(a) maxwell.
(b) gauss.
(c) tesla.
(d) ampere-turn.

10. If a wire coil has 10 turns and carries 500 mA of current, what is the magnetomotive force?
    (a) 5000 At
    (b) 50 At
    (c) 5.0 At
    (d) 0.02 At
11. If a wire coil has 100 turns and carries 1.30 A of current, what is the magnetomotive force?
    (a) 130 Gb
    (b) 76.9 Gb
    (c) 164 Gb
    (d) 61.0 Gb
12. Which of the following can occur during a geomagnetic storm?
    (a) Charged particles stream out from the sun.
    (b) The earth's magnetic field is affected.
    (c) Electrical power transmission is disrupted.
    (d) More than one of the above can occur.
13. An ac electromagnet
    (a) attracts only permanent magnets.
    (b) attracts pure, unmagnetized iron.
    (c) repels all permanent magnets.
    (d) either attracts or repels permanent magnets, depending on the polarity.
14. An advantage of an electromagnet over a permanent magnet is the fact that
    (a) an electromagnet can be switched on and off.
    (b) an electromagnet does not have specific polarity.
    (c) an electromagnet requires no power source.
    (d) permanent magnets must always be cylindrical, but electromagnets can have any shape.
15. A substance with high retentivity
    (a) can make a good ac electromagnet.
    (b) repels both north and south magnetic poles.
    (c) is always a diamagnetic material.
    (d) is well suited to making a permanent magnet.
16. Suppose a relay is connected into a circuit so that a device gets a signal only when the relay coil carries current. The relay is
    (a) an ac relay.
    (b) a dc relay.
    (c) normally closed.
    (d) normally open.

17. A device that repeatedly reverses the polarity of a magnetic field in order to keep a dc motor rotating is known as
    - (a) a solenoid.
    - (b) an armature coil.
    - (c) a commutator.
    - (d) a field coil.

18. A high tape-recorder motor speed is generally used for
    - (a) voice recording and playback.
    - (b) video recording and playback.
    - (c) digital data storage and retrieval.
    - (d) all of the above.

19. An advantage of a magnetic disk, compared with magnetic tape, for data storage and retrieval is that
    - (a) a disk lasts longer.
    - (b) data can be stored and retrieved more quickly with disks than with tapes.
    - (c) disks look better.
    - (d) disks are less susceptible to magnetic fields.

20. A magnetic hard disk is usually part of
    - (a) a computer.
    - (b) a dc motor.
    - (c) a tape recorder.
    - (d) an electromagnet.

Test: Part 1

Do not refer to the text when taking this test. A good score is at least 37 correct. Answers are in the back of the book. It's best to have a friend check your score the first time, so you won't memorize the answers if you want to take the test again.

1. An application in which an analog meter would almost always be preferred over a digital meter is
   (a) the signal-strength indicator in a radio receiver.
   (b) a meter that shows power-supply voltage.
   (c) a utility watt-hour meter.
   (d) a clock.
   (e) a device in which a direct numeric display is wanted.

2. Which of the following statements is false?
   (a) The current in a series dc circuit is divided up among the resistances.
   (b) In a parallel dc circuit, the voltage is the same across each component.
   (c) In a series dc circuit, the sum of the voltages across all the components, going once around a complete circle and taking polarity into account, is zero.
   (d) The net resistance of a parallel set of resistors is less than the value of the smallest resistor.
   (e) The total wattage consumed in a series circuit is the sum of the wattages consumed by each of the components.

3. The ohm is a unit of
   (a) electrical charge quantity.
   (b) the rate at which charge carriers flow.
   (c) opposition to electrical current.
   (d) electrical conductance.
   (e) potential difference.

4. A wiring diagram differs from a schematic diagram in that
   (a) a wiring diagram is less detailed than a schematic diagram.
   (b) a wiring diagram always shows the component values, but a schematic diagram might not.
   (c) a schematic does not show all the interconnections between the components, but a wiring diagram does.
   (d) a schematic diagram shows pictures of components, while a wiring diagram shows the electronic symbols.
   (e) a schematic diagram shows the electronic symbols, while a wiring diagram shows pictures of the components.

5. In which of the following places would you be most likely to find a wirewound resistor?
   (a) A dc circuit location where a large amount of power must be dissipated
   (b) The input circuit of a radio-frequency amplifier
   (c) The output circuit of a radio-frequency amplifier
   (d) In an antenna system, to limit the transmitter power
   (e) Between ground and the chassis of a power supply

6. The number of protons in the nucleus of an element is known as the
   (a) electron number.
   (b) atomic number.
   (c) valence number.
   (d) charge number.
   (e) proton number.

7. A hot-wire ammeter
   (a) can measure ac as well as dc.
   (b) registers current changes very fast.
   (c) can indicate very low voltages.
   (d) measures electrical energy.
   (e) works only when current flows in one direction.

8. Which of the following units indicates the rate at which energy is expended?
   (a) The volt
   (b) The ampere
   (c) The coulomb
   (d) The ampere-hour
   (e) The watt

9. Which of the following correctly states Ohm's Law?
   (a) Volts equal amperes divided by ohms.
   (b) Ohms equal amperes divided by volts.
   (c) Amperes equal ohms divided by volts.

(d) Amperes equal ohms times volts.
(e) Ohms equal volts divided by amperes.

10. The current flowing into a point in a dc circuit is always equal to the current
(a) delivered by the power supply.
(b) through any one of the resistances.
(c) flowing out of that point.
(d) at any other point.
(e) in any single branch of the circuit.

11. A loudness meter in a hi-fi system is generally calibrated in
(a) volts.
(b) amperes.
(c) decibels.
(d) watt-hours.
(e) ohms.

12. An electrically charged atom (either positive or negative) is known as
(a) a molecule.
(b) an isotope.
(c) an ion.
(d) an electron.
(e) a fundamental particle.

13. Suppose a battery delivers 12.0 V to a bulb, and current flowing through the bulb is 3.00 A. The resistance of the bulb is which of the following?
(a) 36.0 Ω
(b) 4.00 Ω
(c) 0.250 Ω
(d) 108 Ω
(e) 0.750 Ω

14. The peak voltage in an ac wave is always
(a) greater than the average voltage.
(b) less than the average voltage.
(c) greater than or equal to the average voltage.
(d) less than or equal to the average voltage.
(e) fluctuating.

15. Suppose a resistor is specified a having a value of 680 Ω, and a tolerance of ±5%. You measure the actual resistance with a precision digital ohmmeter. Which of the following meter readings indicates a reject?
(a) 648 Ω
(b) 712 Ω

(c) 699 Ω
(d) 636 Ω
(e) 707 Ω

16. A primitive device for indicating the presence of an electric current is
   (a) an electrometer.
   (b) a galvanometer.
   (c) a voltmeter.
   (d) a coulometer.
   (e) a wattmeter.
17. A disadvantage of mercury cells is the fact that they
   (a) can adversely affect the environment when discarded.
   (b) supply dangerously high voltage.
   (c) can reverse polarity unexpectedly.
   (d) must be physically larger than other types of cells that have the same current-delivering capacity.
   (e) must be kept right-side up to keep the mercury from spilling out.
18. Suppose a battery supplies 6.0 V to a bulb rated at 12 W. The bulb draws how much current?
   (a) 2.0 A
   (b) 0.5 A
   (c) 72 A
   (d) 40 mA
   (e) 72 mA
19. Which of the following is not a common use for a resistor or set of resistors?
   (a) Biasing for a transistor
   (b) Voltage division
   (c) Current limiting
   (d) As a dummy antenna
   (e) Helping a capacitor to hold its charge for a long time
20. When an electrical charge exists but there is no flow of current, the charge is said to be
   (a) ionizing.
   (b) atomic.
   (c) molecular.
   (d) electronic.
   (e) static.
21. The sum of the voltages, going around a dc circuit, but not including the power supply, has
   (a) an equal value and the same polarity as the supply.
   (b) a value that depends on the ratio of the resistances.
   (c) a different value from, but the same polarity as, the supply.

(d) an equal value as, but the opposite polarity from, the supply.
(e) a different value from, and the opposite polarity from, the supply.

22. A watt-hour meter measures
    (a) voltage.
    (b) current.
    (c) power.
    (d) energy.
    (e) charge.

23. Every chemical element has its own unique type of particle, which is known as its
    (a) neutron.
    (b) electron.
    (c) proton.
    (d) atom.
    (e) isotope.

24. An advantage of a magnetic disk over magnetic tape for data storage is the fact that
    (a) data is too closely packed on the tape.
    (b) the disk is immune to the effects of magnetic fields.
    (c) data storage and retrieval is faster on disk.
    (d) disks store computer data in analog form.
    (e) tapes cannot be used to store digital data.

25. Suppose a 6-V battery is connected across a series combination of resistors. The resistance values are 1.0 Ω, 2.0 Ω, and 3.0 Ω. What is the current through the 2.0-Ω resistor?
    (a) 1.0 A
    (b) 3.0 A
    (c) 12 A
    (d) 24 A
    (e) 72 A

26. A sample of material with resistance so high that it can be considered infinite for most practical purposes is known as
    (a) a semiconductor.
    (b) a paraconductor.
    (c) an insulator.
    (d) a resistor.
    (e) a diamagnetic substance.

27. Primary cells
    (a) can be used over and over.
    (b) have higher voltage than other types of cells.
    (c) all supply exactly 1.500 V.

(d) cannot be recharged.
(e) are made of zinc and carbon.

28. A rheostat
    (a) can be used in high-voltage and/or high-power dc circuits.
    (b) is ideal for tuning a radio receiver.
    (c) is often used as a bleeder resistor.
    (d) is better than a potentiometer for low-power audio.
    (e) offers the advantage of having no inductance.

29. How much dc voltage does a typical dry cell provide?
    (a) 12 V
    (b) 6 V
    (c) 1.5 V
    (d) 117 V
    (e) Any of the above

30. A geomagnetic storm
    (a) causes solar wind.
    (b) causes the earth's magnetic field to disappear.
    (c) can disturb the earth's magnetic field.
    (d) can pollute the earth's atmosphere.
    (e) stabilizes the ac utility grid.

31. An advantage of an alkaline cell over a zinc-carbon cell is the fact that
    (a) the alkaline cell provides more voltage.
    (b) the alkaline cell can be recharged.
    (c) the alkaline cell can deliver useful current at lower temperatures.
    (d) the alkaline cell is far less bulky for the same amount of energy capacity.
    (e) the alkaline cell can produce ac as well as dc.

32. Suppose a battery delivers 12 V across a set of six 4.0-Ω resistors in a series voltage dividing combination. This provides six different voltages, differing by equal increments of which of the following?
    (a) 0.25 V
    (b) 0.33 V
    (c) 1.0 V
    (d) 2.0 V
    (e) 3.0 V

33. A unit of electrical charge quantity is the
    (a) volt.
    (b) ampere.
    (c) watt.

(d) tesla.
(e) coulomb.

34. A unit of conductance is the
    (a) volt per meter.
    (b) ampere per meter.
    (c) anti-ohm.
    (d) siemens.
    (e) ohm per meter.

35. Suppose a 24-V battery is connected across a set of four resistors in parallel. Each resistor has a value of 32 Ω. What is the total power dissipated by the set of resistors?
    (a) 0.19 W
    (b) 3.0 W
    (c) 0.19 kW
    (d) 0.33 W
    (e) 72 W

36. The main difference between a lantern battery and a transistor battery is the fact that
    (a) a lantern battery has higher voltage than a transistor battery.
    (b) a fresh lantern battery has more energy stored in it than a fresh transistor battery.
    (c) a lantern battery cannot be used with electronic devices such as transistor radios, but a transistor battery can.
    (d) a lantern battery can be recharged, but a transistor battery cannot.
    (e) a lantern battery is more compact than a transistor battery.

37. Nickel-based batteries would most likely be found
    (a) in disposable flashlights.
    (b) in large lanterns.
    (c) as car and truck batteries.
    (d) in handheld radio transceivers.
    (e) in electromagnets.

38. A voltmeter should have
    (a) low internal resistance.
    (b) electrostatic plates.
    (c) a sensitive amplifier.
    (d) high internal resistance.
    (e) the highest possible full-scale value.

39. The purpose of a bleeder resistor is to
    (a) provide bias for a transistor.
    (b) serve as a voltage divider.

(c) protect people against the danger of electric shock.
(d) reduce the current in a power supply.
(e) smooth out the ac ripple in a power supply.

40. A dc electromagnet
(a) has constant polarity.
(b) requires an air core.
(c) does not attract or repel a permanent magnet.
(d) has polarity that periodically reverses.
(e) cannot be used to permanently magnetize anything.

41. The rate at which charge carriers flow is measured in
(a) amperes.
(b) coulombs.
(c) volts.
(d) watts.
(e) watt-hours.

42. Suppose a 12-V battery is connected to a set of three resistors in series. The resistance values are 1.0 Ω, 2.0 Ω, and 3.0 Ω. What is the voltage across the 3.0-Ω resistor?
(a) 1.0 V
(b) 2.0 V
(c) 4.0 V
(d) 6.0 V
(e) 12 V

43. Suppose nine 90-Ω resistors are connected in a 3×3 series-parallel network. What is the total (net) resistance of the network?
(a) 10 Ω
(b) 30 Ω
(c) 90 Ω
(d) 270 Ω
(e) 810 Ω

44. A device commonly used for remote switching of high-current circuits is
(a) a solenoid.
(b) an electromagnet.
(c) a potentiometer.
(d) a photovoltaic cell.
(e) a relay.

45. Memory in a nickel-based cell or battery
(a) occurs whenever the battery is discharged.
(b) indicates that the cell or battery is dead.

(c) can usually be remedied by repeated discharging and recharging.
(d) can cause an explosion.
(e) causes a reversal in polarity.

46. Suppose a 100-W bulb burns for 100 hours. It has consumed how many units of energy?
(a) 0.10 kWh
(b) 1.00 kWh
(c) 10.0 kWh
(d) 100 kWh
(e) 1000 kWh

47. A material with high permeability
(a) increases magnetic field quantity.
(b) is necessary if a coil is to produce a magnetic field.
(c) always has high retentivity.
(d) concentrates magnetic lines of flux.
(e) reduces flux density.

48. A chemical compound
(a) consists of two or more atoms.
(b) contains an unusual number of neutrons.
(c) is technically the same as an ion.
(d) has a shortage of electrons.
(e) has an excess of electrons.

49. Suppose a 6.00-V battery is connected to a parallel combination of two resistors whose values are 8.00 Ω and 12.0 Ω. What is the power dissipated in the 8-Ω resistor?
(a) 0.300 W
(b) 0.750 W
(c) 1.25 W
(d) 1.80 W
(e) 4.50 W

50. The main problem with bar-graph meters is the fact that
(a) they are not very sensitive.
(b) they are unstable.
(c) they cannot give very precise readings.
(d) you need special training to read them.
(e) they can display only peak values.

2
PART
Alternating Current

This page intentionally left blank

9
CHAPTER

Alternating-Current Basics

DIRECT CURRENT CAN BE EXPRESSED IN TERMS OF TWO VARIABLES: DIRECTION (POLARITY) AND intensity (amplitude). Alternating current (ac) is a little more complicated. This chapter will acquaint you with some common forms of ac.

Definition of Alternating Current

You have learned that dc has polarity that stays constant over time. Although the amplitude (the number of amperes, volts, or watts) can fluctuate from moment to moment, the charge carriers always flow in the same direction at any point in the circuit.

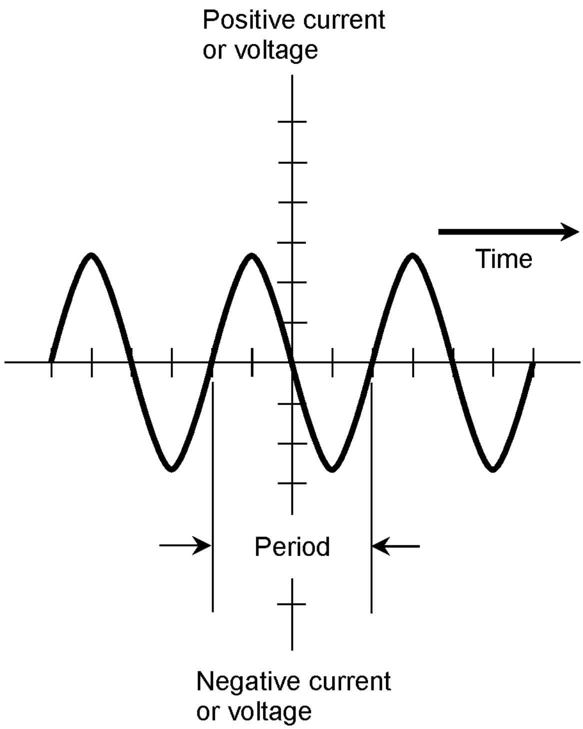
In ac, the polarity reverses at regular intervals. The *instantaneous amplitude* (that is, the amplitude at any given instant in time) of ac usually varies because of the repeated reversal of polarity. But there are certain cases where the amplitude remains constant, even though the polarity keeps reversing.

The rate of change of polarity is the variable that makes ac so much different from dc. The behavior of an ac wave depends largely on this rate: the *frequency.*

Period and Frequency

In a *periodic ac wave*, the kind that is discussed in this chapter (and throughout the rest of this book), the function of *instantaneous amplitude versus time* repeats itself over and over, so that the same pattern recurs indefinitely. The length of time between one repetition of the pattern, or one *cycle*, and the next is called the *period* of the wave. This is illustrated in Fig. 9-1 for a simple ac wave. The period of a wave can, in theory, be anywhere from a minuscule fraction of a second to many centuries. Period, when measured in seconds, is denoted by T.

Originally, ac frequency was specified in *cycles per second* (cps). High frequencies were sometimes given in *kilocycles*, *megacycles*, or *gigacycles*, representing thousands, millions, or billions (thousand-millions) of cycles per second. But nowadays, the unit is known as the *hertz* (Hz). Thus, 1 Hz = 1 cps, 10 Hz = 10 cps, and so on. Higher frequencies are given in *kilohertz* (kHz), *megahertz* (MHz), or *gigahertz* (GHz). The relationships are as follows:



9-1 A sine wave. The period is the length of time it takes for one cycle to be completed.

$$1 \text{ kHz} = 1000 \text{ Hz}$$
$$1 \text{ MHz} = 1000 \text{ kHz} = 1{,}000{,}000 \text{ Hz} = 10^6 \text{ Hz}$$
$$1 \text{ GHz} = 1000 \text{ MHz} = 1{,}000{,}000{,}000 \text{ Hz} = 10^9 \text{ Hz}$$

Sometimes an even bigger unit, the *terahertz* (THz), is used to specify ac frequency. This is a trillion (1,000,000,000,000, or 10^{12}) hertz. Electrical currents generally do not attain such frequencies, although some forms of *electromagnetic radiation* do.

The frequency of an ac wave, denoted f, in hertz is the reciprocal of the period in seconds. Mathematically, these two equations express the relationship:

$$f = 1/T \qquad \text{and} \qquad T = 1/f$$

Some ac waves have only one frequency. These waves are called *pure.* But often, there are components at multiples of the main, or *fundamental*, frequency. There can also be components at odd frequencies. Some ac waves have hundreds, thousands, or even infinitely many different component frequencies.

The Sine Wave

Sometimes, alternating current has a *sine-wave*, or *sinusoidal*, nature. This means that the direction of the current reverses at regular intervals, and that the current-versus-time curve is shaped like the trigonometric *sine function.* The waveform in Fig. 9-1 is a sine wave.

Any ac wave that consists of a single frequency has a perfectly sinusoidal shape. Any perfect sinusoidal ac source has only one component frequency. In practice, a wave might be so close to a sine wave that it looks exactly like the sine function on an oscilloscope, when in reality there are traces of other frequencies present. Imperfections are often too small to see. But pure, single-frequency ac not only looks perfect, but actually is a perfect replication of the trigonometric sine function.

The current at the wall outlets in your house is an almost perfect ac sine wave with a frequency of 60 Hz.

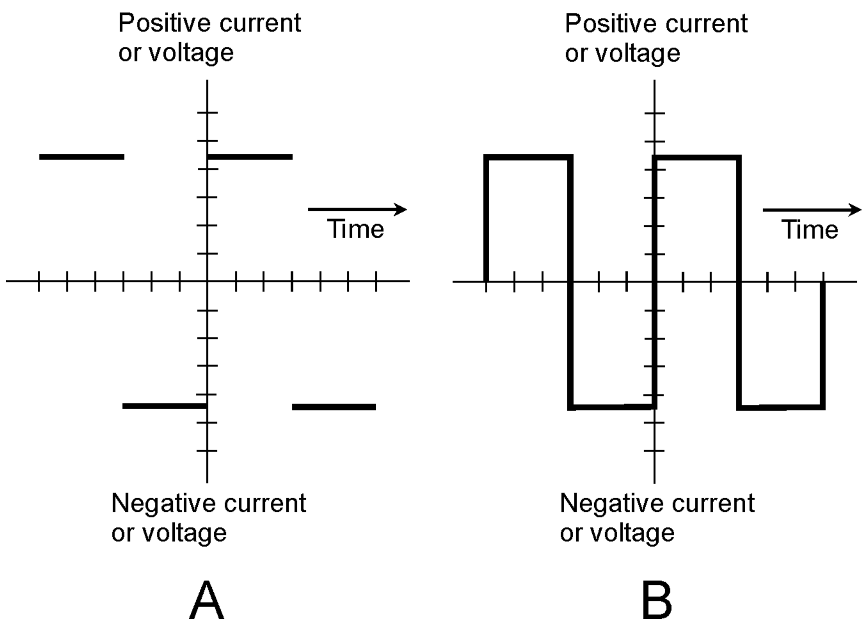
Square Waves

Earlier in this chapter, it was said that there can be an ac wave whose instantaneous amplitude remains constant, even though the polarity reverses. Does this seem counterintuitive? Think some more! A *square wave* is such a wave.

On an oscilloscope, a square wave looks like a pair of parallel, dashed lines, one with positive polarity and the other with negative polarity (Fig. 9-2A). The oscilloscope shows a graph of voltage on the vertical scale and time on the horizontal scale. The transitions between negative and positive for a theoretically perfect square wave would not show up on the oscilloscope, because they would be instantaneous. But in practice, the transitions can often be seen as vertical lines (Fig. 9-2B).

True square waves have equal negative and positive peaks. Thus, the absolute amplitude of the wave is constant. Half of the time it's $+x$, and the other half of the time it's $-x$ (where x can be expressed in volts, amperes, or watts).

Some squared-off waves are lopsided; the negative and positive amplitudes are not the same. Still others remain at positive polarity longer than they remain at negative polarity (or vice versa). These are examples of *asymmetrical square waves*, more properly called *rectangular waves.*



9-2 At A, a perfect square wave; the transitions are instantaneous and therefore do not show up on the graph. At B, the more common rendition of a square wave, showing the transitions as vertical lines.

Sawtooth Waves

Some ac waves rise and/or fall in straight, sloping lines as seen on an oscilloscope screen. The slope of the line indicates how fast the magnitude is changing. Such waves are called *sawtooth waves* because of their appearance. Sawtooth waves are generated by certain electronic test devices. They can also be generated by electronic sound synthesizers.

Fast Rise, Slow Decay

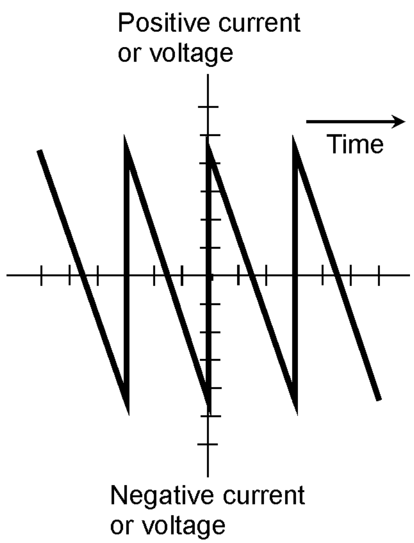
Figure 9-3 shows a sawtooth wave in which the positive-going slope (called the *rise*) is extremely steep, as with a square wave, but the negative-going slope (called the *decay*) is not so steep. The period of the wave is the time between points at identical positions on two successive pulses.

Slow Rise, Fast Decay

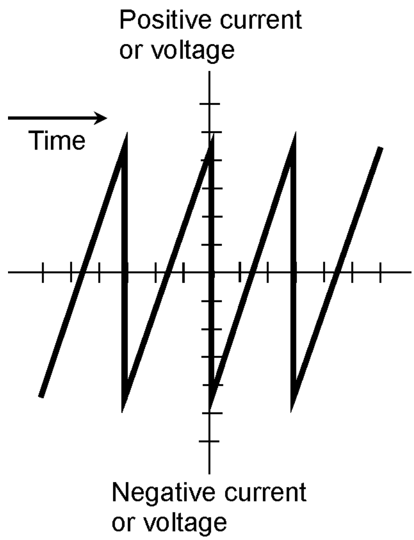
Another form of sawtooth wave is just the opposite, with a defined, finite rise and an instantaneous decay. This type of wave is often called a *ramp* because it looks like an incline going upward (Fig. 9-4). This waveshape is useful for scanning in television sets and oscilloscopes. It tells the electron beam to move, or *trace*, at constant speed from left to right across the screen during the rise. Then it retraces, or brings the electron beam back, instantaneously during the decay so the beam can trace across the screen again.

Variable Rise and Decay

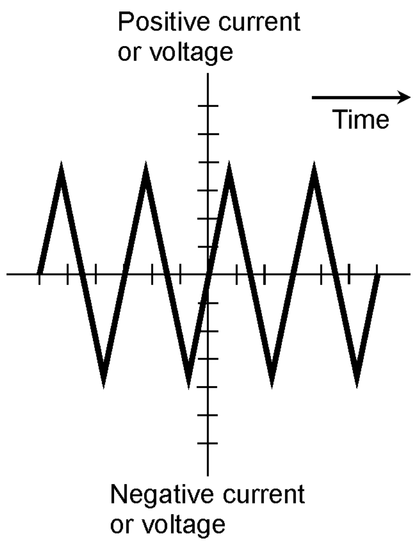
Sawtooth waves can have rise and decay slopes in an infinite number of different combinations. One common example is shown in Fig. 9-5. In this case, the rise and the decay are both finite and equal. This is known as a *triangular wave.*



9-3 A sawtooth wave with a fast rise and a slow decay.



9-4 A sawtooth wave with a slow rise and a fast decay.

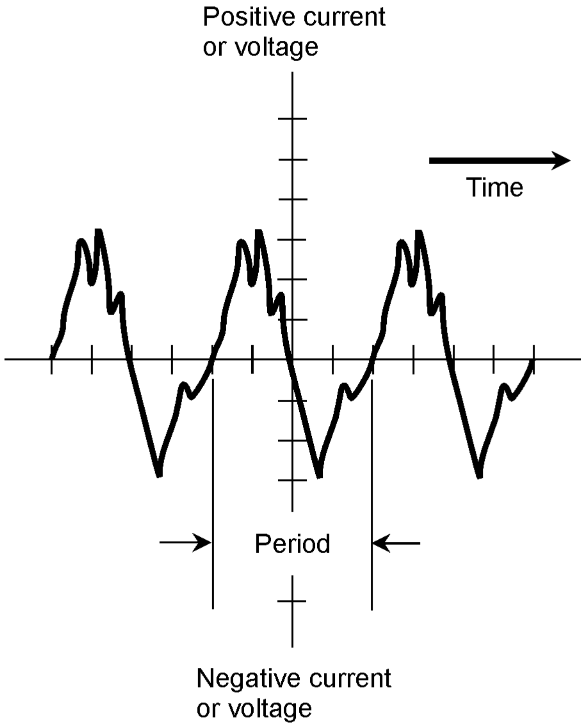


9-5 A triangular wave with rise and decay rates that are the same.

Complex and Irregular Waveforms

As long as a wave has a definite period, and as long as the polarity keeps switching back and forth between positive and negative, it is ac, no matter how complicated the actual shape of the waveform. Figure 9-6 shows an example of a complex ac wave. There is a definable period, and therefore a definable frequency. The period is the time between two points on succeeding wave repetitions.

With some waves, it can be difficult or almost impossible to ascertain the period. This is because the wave has two or more components that are of nearly the same amplitude. When this hap-



9-6 An irregular waveform.

pens, the *frequency spectrum* of the wave is multifaceted. That means the wave energy is split up more or less equally among multiple frequencies.

Frequency Spectrum

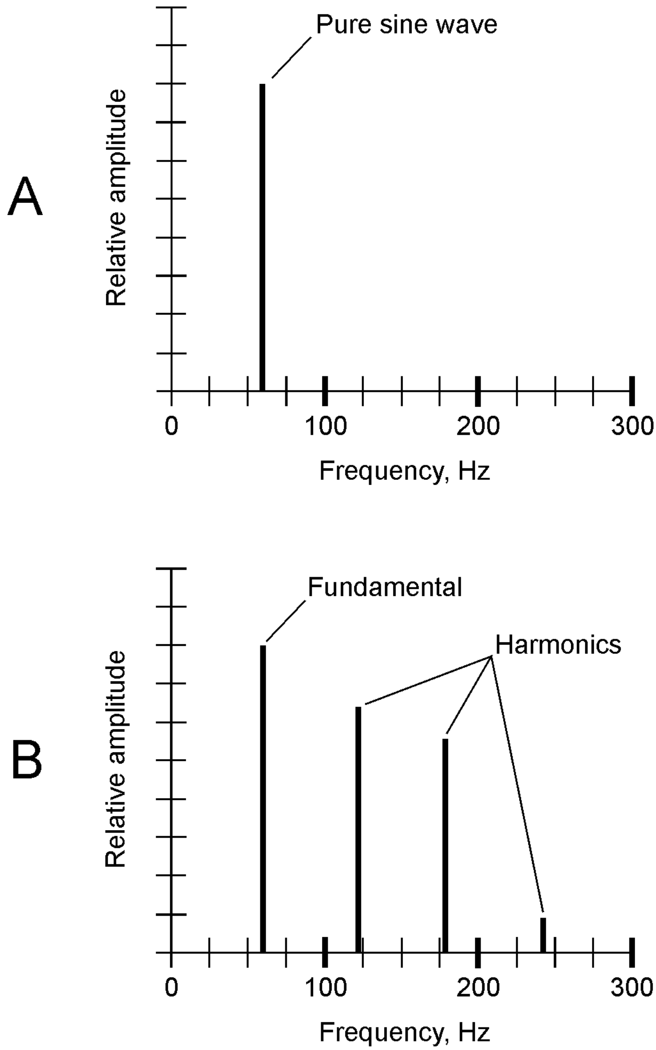
An oscilloscope shows a graph of amplitude as a function of time. Because time is on the horizontal axis and represents the *independent variable* or *domain* of the function, the oscilloscope is said to be a *time-domain* instrument. But suppose you want to see the amplitude of a complex signal as a function of frequency, rather than as a function of time? This can be done with a *spectrum analyzer.* It is a *frequency-domain* instrument. Its horizontal axis shows frequency as the independent variable, ranging from some adjustable minimum frequency (at the extreme left) to some adjustable maximum frequency (at the extreme right).

An ac sine wave, as displayed on a spectrum analyzer, appears as a single *pip*, or vertical line (Fig. 9-7A). This means that all of the energy in the wave is concentrated at one frequency. But many, if not most, ac waves contain *harmonic* energy along with energy at the fundamental frequency. A harmonic frequency is a whole-number multiple of the fundamental frequency. For example, if 60 Hz is the fundamental frequency, then harmonics can exist at 120 Hz, 180 Hz, 240 Hz, and so on. The 120-Hz wave is the *second harmonic;* the 180-Hz wave is the *third harmonic;* the 240-Hz wave is the *fourth harmonic;* and so on.

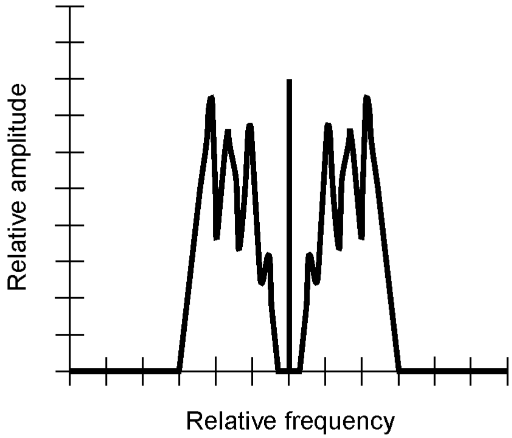
In general, if a wave has a frequency equal to *n* times the fundamental (where *n* is some whole number), then that wave is called the n*th harmonic.* In Fig. 9-7B, a wave is shown along with several harmonics, as it would look on the display screen of a spectrum analyzer.

Square waves and sawtooth waves contain harmonic energy in addition to energy at the fundamental frequency. Other waves can get more complicated. The exact shape of a wave depends on the amount of energy in the harmonics, and the way in which this energy is distributed among them.

Irregular waves can have any imaginable frequency distribution. Figure 9-8 shows an example. This is a spectral (frequency-domain) display of an *amplitude-modulated* (AM) voice radio signal. Much of the energy is concentrated at the center of the pattern, at the frequency shown by the vertical line. That is the *carrier frequency.* There is also plenty of energy near, but not exactly at, the carrier frequency. That's the part of the signal that contains the voice.



9-7 At A, a spectral diagram of a pure, 60-Hz sine wave. At B, a spectral diagram of a 60-Hz wave with three harmonics.



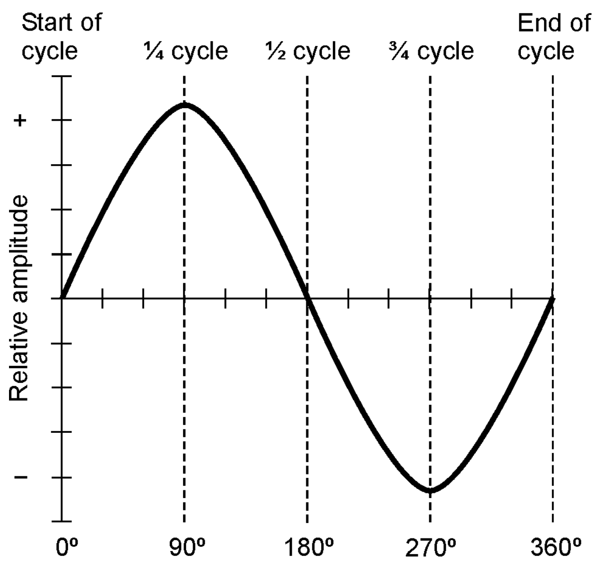
9-8 A spectral diagram of a modulated radio signal.

Fractions of a Cycle

Engineers break the ac cycle down into small parts for analysis and reference. One complete cycle can be compared to a single revolution around a circle.

Degrees

One method of specifying the phase of an ac cycle is to divide it into 360 equal parts, called *degrees* or *degrees of phase*, symbolized by a superscript, lowercase letter *o* (°). The value 0° is assigned to the point in the cycle where the magnitude is zero and positive-going. The same point on the next cycle is given the value 360°. The point one-fourth of the way through the cycle is 90°; the point halfway through the cycle is 180°; the point three-fourths of the way through the cycle is 270°. This is illustrated in Fig. 9-9. Degrees of phase are used mainly by engineers and technicians.



9-9 A cycle is divided into 360 equal parts, called *degrees.*

Radians

The other method of specifying phase is to divide the cycle into 2π equal parts, where π (pi) is a geometric constant equal to the number of diameters of any circle that can be laid end to end around the circumference of that circle. This constant is approximately equal to 3.14159. A *radian* (rad) of phase is thus equal to about 57.3°. Sometimes, the frequency of an ac wave is measured in *radians per second* (rad/s) rather than in hertz. Because there are about 6.28 radians in a complete cycle of 360°, the *angular frequency* of a wave, in radians per second, is equal to about 6.28 times the frequency in hertz. Radians of phase are used mainly by physicists.

Phase Difference

Even if two ac waves have exactly the same frequency, they can have different effects because they are out of sync with each other. This is especially true when ac waves are added together to produce a third, or *composite*, wave.

If two pure ac sine waves have identical frequencies and identical amplitudes but differ in phase by 180° (a half cycle), they cancel each other out, and the composite wave is zero; it ceases to exist! If the two waves are exactly in phase, the composite wave has the same frequency, but twice the amplitude, of either signal alone.

If two pure ac sine waves have the same frequency but different amplitudes, and if they differ in phase by 180°, the composite signal has the same frequency as the originals, and an amplitude equal to the difference between the two. If two such waves are exactly in phase, the composite has the same frequency as the originals, and an amplitude equal to the sum of the two.

If two pure ac sine waves have the same frequency but differ in phase by some odd amount such as 75° or 110°, the resulting signal has the same frequency, but does not have the same waveshape as either of the original signals. The variety of such cases is infinite.

Household electricity from 117-V wall outlets consists of a 60-Hz sine wave with only one phase component. But the energy is transmitted over long distances in three phases, each differing by 120° or one-third of a cycle. This is what is meant by *three-phase ac.* Each of the three ac waves carries one-third of the total power in a utility transmission line.

Expressions of Amplitude

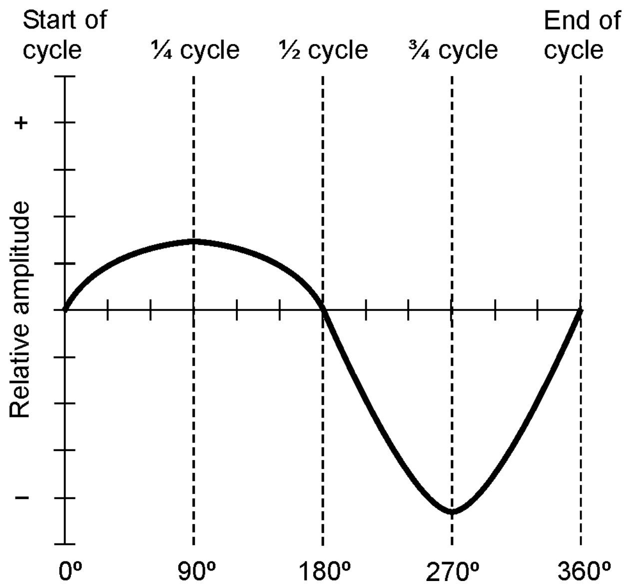
Amplitude is also called *magnitude*, *level*, *strength*, or *intensity.* Depending on the quantity being measured, the amplitude of an ac wave can be specified in amperes (for current), volts (for voltage), or watts (for power). In addition to this, there are several different ways in which amplitude can be expressed.

Instantaneous Amplitude

The *instantaneous amplitude* of an ac wave is the amplitude at some precise moment, or instant, in time. This constantly changes. The manner in which it varies depends on the waveform. Instantaneous amplitudes are represented by individual points on the wave curves.

Peak Amplitude

The *peak* (pk) *amplitude* of an ac wave is the maximum extent, either positive or negative, that the instantaneous amplitude attains. In many situations, the positive and negative peak amplitudes of

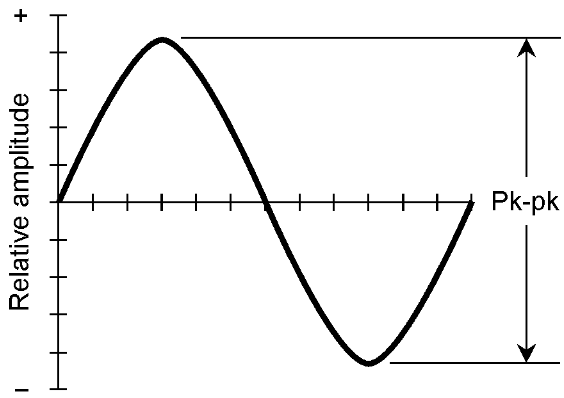


9-10 A wave with unequal positive and negative peak amplitudes.

an ac wave are the same. But sometimes they differ. Figure 9-9 is an example of a wave in which the positive peak amplitude is the same as the negative peak amplitude. Figure 9-10 is an illustration of a wave that has different positive and negative peak amplitudes.

Peak-to-Peak Amplitude

The *peak-to-peak* (pk-pk) *amplitude* of a wave is the net difference between the positive peak amplitude and the negative peak amplitude (Fig. 9-11). The peak-to-peak amplitude is equal to the positive peak amplitude plus the negative peak amplitude. When the positive and negative peak amplitudes of an ac wave are equal, the peak-to-peak amplitude is exactly twice the peak amplitude.



9-11 Peak-to-peak (pk-pk) amplitude of a sine wave.

Root-Mean-Square Amplitude

Often, it is necessary to express the *effective amplitude* of an ac wave. This is the voltage, current, or power that a dc source would have to produce in order to have the same general effect as a given ac wave. When you say a wall outlet provides 117 V, you mean 117 effective volts. This is not the same as the peak or peak-to-peak voltage.

The most common expression for effective ac intensity is called the *root-mean-square* (rms) *amplitude.* The terminology reflects the fact that the ac wave is mathematically operated on by taking the square root of the mean (average) of the square of all its instantaneous amplitudes.

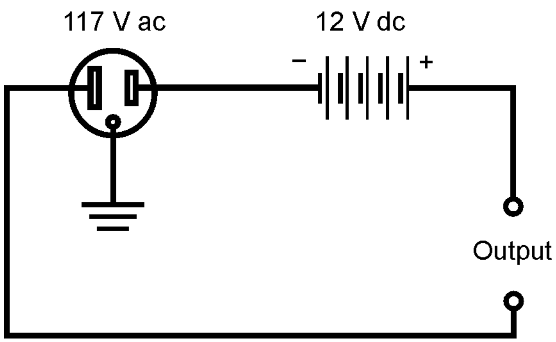
In the case of a perfect ac sine wave, the rms value is equal to 0.707 times the peak value, or 0.354 times the peak-to-peak value. Conversely, the peak value is 1.414 times the rms value, and the peak-to-peak value is 2.828 times the rms value. The rms amplitude is often specified when talking about utility ac, radio-frequency (RF) ac, and audio-frequency (AF) ac.

For a perfect square wave, the rms value is the same as the peak value, and half the peak-to-peak value. For sawtooth and irregular waves, the relationship between the rms value and the peak value depends on the exact shape of the wave. But the rms value is never greater than the peak value for any type of ac wave.

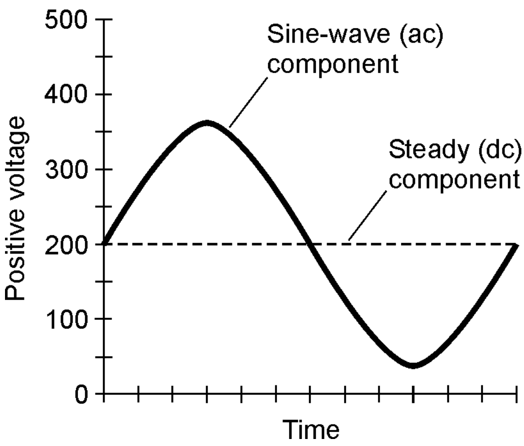
Superimposed DC

Sometimes a wave has components of both ac and dc. The simplest example of an ac/dc combination is illustrated by the connection of a dc voltage source, such as a battery, in series with an ac voltage source, like the utility mains. An example is shown in the schematic diagram of Fig. 9-12. Imagine connecting a 12-V automotive battery in series with the wall outlet. (Do not try this experiment in real life!) When this is done, the ac wave is displaced either positively or negatively by 12 V, depending on the polarity of the battery. This results in a sine wave at the output, but one peak is 24 V (twice the battery voltage) more than the other.

Any ac wave can have dc components along with it. If the dc component exceeds the peak value of the ac wave, then fluctuating, or pulsating, dc will result. This would happen, for example, if a 200-V dc source were connected in series with the output of a common utility ac outlet, which has peak voltages of approximately ± 165 V. Pulsating dc would appear, with an average value of 200 V but with instantaneous values much higher and lower. The waveshape in this case is shown in Fig. 9-13.



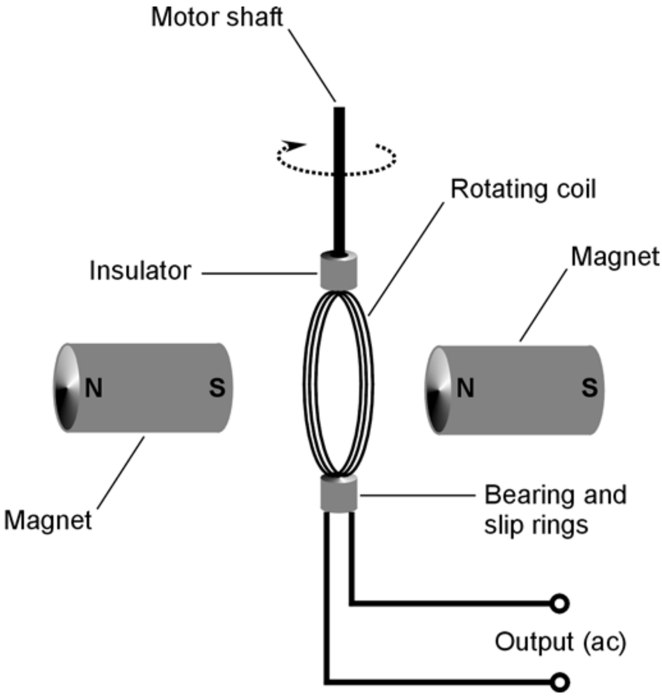
9-12 Connection of a dc source in series with an ac source.



9-13 Waveform resulting from a 117-V ac sine-wave source connected in series with a +200-V dc source.

The Generator

Alternating current can be generated by a rotating coil of wire inside a powerful magnet, as shown in Fig. 9-14. An ac voltage appears between the ends of the wire coil. The ac voltage that a generator can produce depends on the strength of the magnet, the number of turns in the wire coil, and the speed at which the magnet or coil rotates. The ac frequency depends only on the speed of rota-



9-14 A functional diagram of an ac generator.

tion. Normally, for utility ac, this speed is 3600 revolutions per minute (rpm), or 60 complete revolutions per second (rps), so the ac output frequency is 60 Hz.

When a load, such as a light bulb or heater, is connected to an ac generator, it becomes more difficult, mechanically, to turn the generator shaft, compared to when there is nothing connected to the output. As the amount of electrical power demanded from a generator increases, so does the mechanical power required to drive it. This is why it is impossible to connect a generator to a stationary bicycle and pedal an entire city into electrification. There's no way to get something for nothing. The electrical power that comes out of a generator can never be more than the mechanical power driving it. In fact, there is always some energy lost, mainly as heat in the generator. Your legs might generate enough power to run a small radio or television set, but nowhere near enough to provide electricity for a household.

The *efficiency* of a generator is the ratio of the electrical power output to the mechanical driving power, both measured in the same units (such as watts or kilowatts), multiplied by 100 to get a percentage. No generator is 100 percent efficient, but a good one can come fairly close.

At power plants, generators are driven by massive turbines. The turbines are turned by various natural sources of energy such as moving water, steam heated by combustion of fossil fuels, or steam taken directly from deep inside the earth. These energy sources can provide tremendous mechanical power, and this is why power plants can produce megawatts of electrical power.

Why Alternating and Not Direct?

Do you wonder why ac is used at all? Isn't it a lot more complicated than dc? Well, ac may be more complicated in theory, but in practice it is a lot simpler to use when it is necessary to provide electricity to a large number of people.

Alternating current lends itself well to being transformed to lower or higher voltages, according to the needs of electrical apparatus. It is not so easy to change dc voltages. Electrochemical cells produce dc directly, but they are impractical for the needs of large populations. Serving millions of consumers requires the immense power of falling or flowing water, the ocean tides, wind, fossil fuels, controlled nuclear reactions, or geothermal heat. All of these energy sources can be used to drive turbines that turn ac generators.

Technology is advancing in the realm of solar-electric energy; someday a significant part of our electricity might come from photovoltaic power plants. These would generate dc. High voltages could be attained by connecting giant arrays of solar panels in series. But there would be a problem transforming this voltage down to manageable levels for consumer use.

Thomas Edison is said to have favored dc over ac for electrical power transmission in the early days, as the electric utilities were first being devised and constructed. His colleagues argued that ac would work better. But perhaps Edison knew something that his contemporaries did not. There is one advantage to dc in utility applications, and it involves the transmission of energy over great distances using wires. Direct currents, at extremely high voltages, are transported more efficiently than alternating currents. The wire has less effective resistance with dc than with ac, and there is less energy lost in the magnetic fields around the wires. Direct-current *high-tension* transmission lines are being considered for future use. Right now, the main problem is expense. Sophisticated power-conversion equipment is needed. If the cost can be brought within reason, Edison will be vindicated.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct. Answers are in the back of the book.

1. Which of the following can vary with ac, but never with dc?
   (a) Power
   (b) Voltage
   (c) Frequency
   (d) Amplitude

2. The length of time between a point in one cycle and the same point in the next cycle of an ac wave is the
   (a) frequency.
   (b) magnitude.
   (c) period.
   (d) polarity.

3. On a spectrum analyzer, an ac signal having only one frequency component looks like
   (a) a single pip.
   (b) a sine wave.
   (c) a square wave.
   (d) a sawtooth wave.

4. The period of an ac wave, in seconds, is
   (a) the same as the frequency in hertz.
   (b) not related to the frequency in any way.
   (c) equal to 1 divided by the frequency in hertz.
   (d) equal to the peak amplitude in volts divided by the frequency in hertz.

5. The sixth harmonic of an ac wave whose period is 1.000 millisecond (1.000 ms) has a frequency of
   (a) 0.006 Hz.
   (b) 167.0 Hz.
   (c) 7.000 kHz.
   (d) 6.000 kHz.

6. A degree of phase represents
   (a) 6.28 cycles.
   (b) 57.3 cycles.
   (c) $\frac{1}{60}$ of a cycle.
   (d) $\frac{1}{360}$ of a cycle.

7. Suppose that two ac waves have the same frequency but differ in phase by exactly 1⁄20 of a cycle. What is the phase difference between these two waves?
   (a) 18°
   (b) 20°
   (c) 36°
   (d) 5.73°

8. Suppose an ac signal has a frequency of 1770 Hz. What is its angular frequency?
   (a) 1770 rad/s
   (b) 11,120 rad/s
   (c) 282 rad/s
   (d) Impossible to determine from the data given

9. A triangular wave exhibits
   (a) an instantaneous rise and a defined decay.
   (b) a defined rise and an instantaneous decay.
   (c) a defined rise and a defined decay, and the two are equal.
   (d) an instantaneous rise and an instantaneous decay.

10. Three-phase ac
   (a) has sawtooth waves that add together in phase.
   (b) consists of three sine waves in different phases.
   (c) is a sine wave with exactly three harmonics.
   (d) is of interest only to physicists.

11. If two perfect sine waves have the same frequency and the same amplitude, but are in opposite phase, the composite wave
   (a) has twice the amplitude of either input wave alone.
   (b) has half the amplitude of either input wave alone.
   (c) is complex, but has the same frequency as the originals.
   (d) has zero amplitude (that is, it does not exist), because the two input waves cancel each other out.

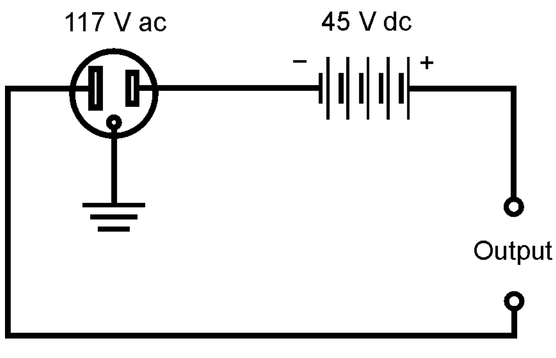
12. If two perfect sine waves have the same frequency and the same phase, the composite wave
   (a) is a sine wave with an amplitude equal to the difference between the amplitudes of the two input waves.
   (b) is a sine wave with an amplitude equal to the sum of the amplitudes of the two original waves.
   (c) is not a sine wave, but has the same frequency as the two input waves.
   (d) has zero amplitude (that is, it does not exist), because the two input waves cancel each other out.

13. In a 117-V rms utility circuit, the positive peak voltage is approximately
    (a) +82.7 V.
    (b) +165 V.
    (c) +234 V.
    (d) +331 V.

14. In a 117-V rms utility circuit, the peak-to-peak voltage is approximately
    (a) 82.7 V.
    (b) 165 V.
    (c) 234 V.
    (d) 331 V.

15. In a perfect sine wave, the peak-to-peak amplitude is equal to
    (a) half the peak amplitude.
    (b) the peak amplitude.
    (c) 1.414 times the peak amplitude.
    (d) twice the peak amplitude.

16. If a 45-V dc battery is connected in series with the 117-V rms utility mains as shown in Fig. 9-15, the peak voltages will be approximately
    (a) +210 V and −120 V.
    (b) +162 V and −72 V.
    (c) +396 V and −286 V.
    (d) +117 V and −117V.



9-15 Illustration for Quiz Question 16.

17. In the situation described in question 16 and illustrated in Fig. 9-15, the peak-to-peak voltage will be approximately
    (a) 117 V.
    (b) 210 V.
    (c) 331 V.
    (d) 396 V.

18. Which one of the following does *not* affect the power output available from a particular ac generator?
    (a) The strength of the magnet
    (b) The number of turns in the coil
    (c) The type of natural energy source used
    (d) The speed of rotation of the coil or magnet

19. If a 175-V dc source were connected in series with the utility mains from a standard wall outlet, the result would be
    (a) smooth dc at a constant voltage.
    (b) pure ac with equal peak voltages.
    (c) ac with one peak voltage greater than the other.
    (d) fluctuating dc.

20. An advantage of ac over dc in utility applications is the fact that
    (a) ac is easier to transform from one voltage to another.
    (b) ac is transmitted with lower loss in wires.
    (c) ac can be easily obtained from dc generators.
    (d) ac can be generated with less-dangerous by-products.

10
CHAPTER

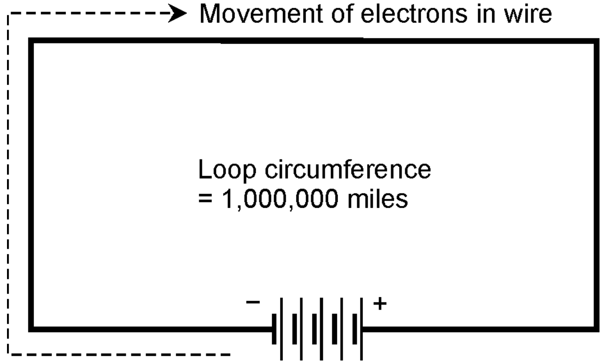
Inductance

IN THIS CHAPTER, YOU'LL LEARN ABOUT ELECTRICAL COMPONENTS THAT OPPOSE THE FLOW OF AC BY temporarily storing energy as magnetic fields. These devices are called *inductors*, and their action is known as *inductance.* Inductors often, but not always, consist of wire coils. Sometimes a length of wire, or a pair of wires, is used as an inductor.

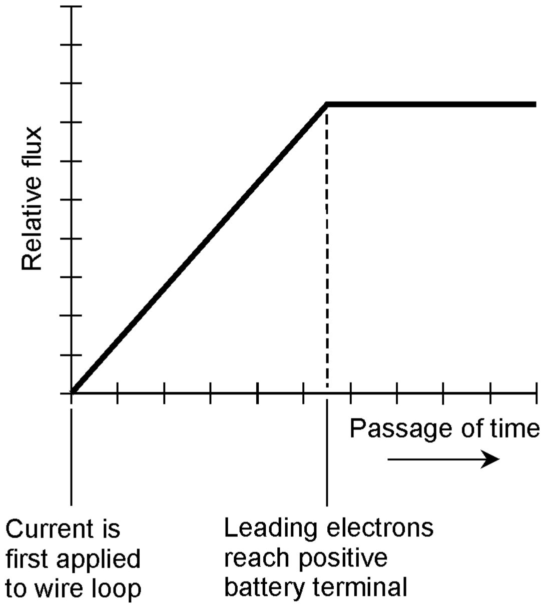
The Property of Inductance

Suppose you have a wire 1 million miles long (about 1.6 million kilometers). Imagine that you make this wire into a huge loop, and connect its ends to the terminals of a battery (Fig. 10-1). An electrical current will flow through the loop of wire, but this is only part of the picture.

If the wire was short, the current would begin to flow immediately, and it would attain a level limited by the resistance in the wire and in the battery. But because the wire is extremely long, it takes a while for the electrons from the negative terminal to work their way around the loop to the positive terminal. It will take a little time for the current to build up to its maximum level.



10-1 A huge, imaginary loop of wire can be used to illustrate the principle of inductance.



10-2 Relative magnetic flux in and around a huge loop of wire connected to a current source, as a function of time.

The magnetic field produced by the loop will be small during the first few moments when current flows in only part of the loop. The magnetic field will build up as the electrons get around the loop. Once a steady current is flowing around the entire loop, the magnetic field will have reached its maximum quantity and will level off (see Fig. 10-2). A certain amount of energy is stored in this magnetic field. The amount of stored energy depends on the *inductance* of the loop, which is a function of its overall size. Inductance, as a property or as a mathematical variable, is symbolized by an italicized, uppercase letter *L*. The loop constitutes an *inductor*, the symbol for which is an uppercase, nonitalicized letter L.

Practical Inductors

It is impractical to make wire loops 1 million miles in circumference. But lengths of wire can be coiled up. When this is done, the magnetic flux is increased for a given length of wire compared with the flux produced by a single-turn loop.

The magnetic flux density inside a coil is multiplied when a ferromagnetic core is placed within it. The increase in flux density has the effect of increasing the inductance, too, so *L* is many times greater with a ferromagnetic core than with an air core or a nonmagnetic core such as plastic or wood. The current that an inductor can handle depends on the diameter (gauge) of the wire. But the value of *L* is a function of the number of turns in the coil, the diameter of the coil itself, and the overall shape of the coil.

In general, the inductance of a coil is directly proportional to the number of turns of wire. Inductance is directly proportional to the diameter of the coil. The length of a coil, given a certain number of turns and a certain diameter, has an effect as well. If a coil having a certain number of turns and a certain diameter is "stretched out," its inductance decreases. Conversely, if it is "squashed up," its inductance increases.

The Unit of Inductance

When a battery is first connected across an inductor, the current builds up at a rate that depends on the inductance. The greater the inductance, the slower the rate of current buildup for a given battery voltage. The unit of inductance is an expression of the ratio between the rate of current buildup and the voltage across an inductor. An inductance of 1 *henry* (1 H) represents a potential difference of 1 volt (1 V) across an inductor within which the current is changing at the rate of 1 ampere per second (1 A/s).

The henry is a huge unit of inductance. You won't often see an inductor this large, although some power-supply filter chokes have inductances up to several henrys. Usually, inductances are expressed in *millihenrys* (mH), *microhenrys* (μH), or *nanohenrys* (nH). You should know your prefix multipliers by now, but in case you've forgotten:

$$1 \text{ mH} = 0.001 \text{ H} = 10^{-3} \text{ H}$$
$$1 \text{ μH} = 0.001 \text{ mH} = 10^{-6} \text{ H}$$
$$1 \text{ nH} = 0.001 \text{ μH} = 10^{-9} \text{ H}$$

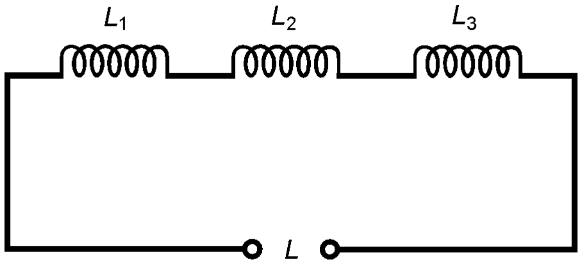
Small coils with few turns of wire produce small inductances, in which the current changes quickly and the induced voltages are small. Large coils with ferromagnetic cores, and having many turns of wire, have high inductances in which the current changes slowly and the induced voltages are large. The current from a battery, building up or dying down through a high-*L* coil, can give rise to a deadly potential difference between the end terminals of the coil—many times the voltage of the battery itself. This is how spark coils work in internal combustion engines. Be careful around them!

Inductors in Series

When the magnetic fields around inductors do not interact, inductances in series add like resistances in series. The total value is the sum of the individual values. It's important to be sure that you are using the same size units for all the inductors when you add their values. After that, you can convert the result to any inductance unit you want.

Problem 10-1

Suppose three 40.0-μH inductors are connected in series, and there is no interaction, or *mutual inductance*, among them (Fig. 10-3). What is the total inductance?



10-3 Inductances in series simply add up, as long as the inductors do not interact.

Add up the values. Call the inductances of the individual components L_1, L_2, and L_3, and the total inductance L. Then $L = L_1 + L_2 + L_3 = 40.0 + 40.0 + 40.0 = 120$ μH.

Problem 10-2

Imagine three inductors, with no mutual inductance, with values of 20.0 mH, 55.0 μH, and 400 nH. What is the total inductance, in millihenrys, of these components if they are connected in series as shown in Fig. 10-3?

First, convert all the inductances to the same units. Microhenrys are a good choice because that unit makes the calculation process the least messy. Call L_1 = 20.0 mH = 20,000 μH, L_2 = 55.0 μH, and L_3 = 400 nH = 0.400 μH. The total inductance is therefore L = 20,000 + 55.0 + 0.400 = 20,055.4 μH. This is 20.1 mH after converting and rounding off.

Inductors in Parallel

If there is no mutual inductance among two or more parallel-connected inductors, their values add up like the values of resistors in parallel. Suppose you have inductances $L_1, L_2, L_3, \ldots, L_n$ all connected in parallel. Then you can find the reciprocal of the total inductance, $1/L$, using the following formula:

$$1/L = 1/L_1 + 1/L_2 + 1/L_3 + \ldots + 1/L_n$$

The total inductance, L, is found by taking the reciprocal of the number you get for $1/L$. Again, as with inductances in series, it's important to remember that all the units have to agree during the calculation process. Once you have completed the calculation, you can convert the result to any inductance unit.

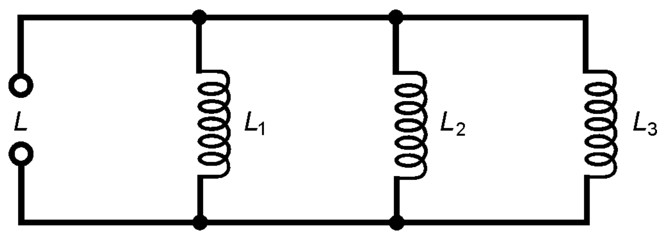
Problem 10-3

Suppose there are three inductors, each with a value of 40 μH, connected in parallel with no mutual inductance, as shown in Fig. 10-4. What is the net inductance of the combination?

Let's call the inductances L_1 = 40 μH, L_2 = 40 μH, and L_3 = 40 μH. Use the preceding formula to obtain $1/L = 1/40 + 1/40 + 1/40 = 3/40 = 0.075$. Then $L = 1/0.075 = 13.333$ μH. This should be rounded off to 13 μH, because the original inductances are specified to only two significant digits.

Problem 10-4

Imagine four inductors in parallel, with no mutual inductance and values of L_1 = 75.0 mH, L_2 = 40.0 mH, L_3 = 333 μH, and L_4 = 7.00 H. What is the net inductance of this combination?



10-4 Inductances in parallel.

You can use henrys, millihenrys, or microhenrys as the standard units in this problem. Suppose you decide to use henrys. Then $L_1 = 0.0750$ H, $L_2 = 0.0400$ H, $L_3 = 0.000333$ H, and $L_4 = 7.00$ H. Use the preceding formula to obtain $1/L = 13.33 + 25.0 + 3003 + 0.143 = 3041.473$. The reciprocal of this is the inductance $L = 0.00032879$ H $= 328.79$ μH. This should be rounded off to 329 μH. This is only a little less than the value of the 333 μH inductor alone.

If there are several inductors in parallel, and one of them has a value that is much smaller than the values of all the others, then the total inductance is a little smaller than the value of the smallest inductor.

Interaction among Inductors

In real-world circuits, there is almost always some mutual inductance between or among solenoidal coils. The magnetic fields extend significantly outside such coils, and mutual effects are difficult to avoid or eliminate. The same is true between and among lengths of wire, especially at high ac frequencies. Sometimes, mutual inductance has no detrimental effect, but in some situations it is not wanted. Mutual inductance can be minimized by using *shielded* wires and *toroidal* inductors. The most common shielded wire is *coaxial cable.* Toroidal inductors are discussed later in this chapter.

Coefficient of Coupling

The *coefficient of coupling*, symbolized k, is an expression of the extent to which two inductors interact. It is specified as a number ranging from 0 (no interaction) to 1 (the maximum possible interaction). Two coils separated by a sheet of solid iron, or by a great distance, have a coefficient of coupling of zero ($k = 0$); two coils wound on the same form, one right over the other, have the maximum possible coefficient of coupling ($k = 1$). Sometimes, the coefficient of coupling is multiplied by 100 and expressed as a percentage from 0 to 100 percent.

Mutual Inductance

The *mutual inductance* between two inductors is symbolized M, and is expressed in the same units as inductance: henrys, millihenrys, microhenrys, or nanohenrys. The value of M is a function of the values of the inductors, and also of the coefficient of coupling.

In the case of two inductors having values of L_1 and L_2 (both expressed in the same size units), and with a coefficient of coupling equal to k, the mutual inductance M is found by multiplying the inductance values, taking the square root of the result, and then multiplying by k. Mathematically:

$$M = k(L_1 L_2)^{1/2}$$

where the ½ power represents the square root. The value of M thus obtained will be in the same size unit as the values of the inductance you input to the equation.

Effects of Mutual Inductance

Mutual inductance can either increase or decrease the net inductance of a pair of series-connected coils, compared with the condition of zero mutual inductance. The magnetic fields around the coils either reinforce each other or oppose each other, depending on the phase relationship of the ac applied to them. If the two ac waves (and thus the magnetic fields they produce) are in phase, the inductance is increased compared with the condition of zero mutual inductance. If the two waves are in opposing phase, the net inductance is decreased relative to the condition of zero mutual inductance.

When two inductors are connected in series and there is *reinforcing* mutual inductance between them, the total inductance L is given by the following formula:

$$L = L_1 + L_2 + 2M$$

where L_1 and L_2 are the inductances, and M is the mutual inductance. All inductances must be expressed in the same size units.

When two inductors are connected in series and the mutual inductance is *opposing*, the total inductance L is given by this formula:

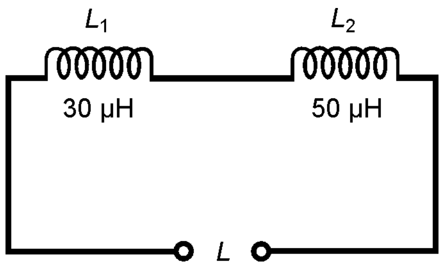
$$L = L_1 + L_2 - 2M$$

where, again, L_1 and L_2 are the values of the individual inductors.

It is possible for mutual inductance to increase the total series inductance of a pair of coils by as much as a factor of 2, if the coupling is total and if the flux reinforces. Conversely, it is possible for the inductances of two coils to completely cancel each other. If two equal-valued inductors are connected in series so their fluxes oppose (or *buck* each other) and $k = 1$, the result is theoretically zero inductance.

Problem 10-5

Suppose two coils, having inductances of 30 μH and 50 μH, are connected in series so that their fields reinforce, as shown in Fig. 10-5. Suppose that the coefficient of coupling is 0.500. What is the total inductance of the combination?

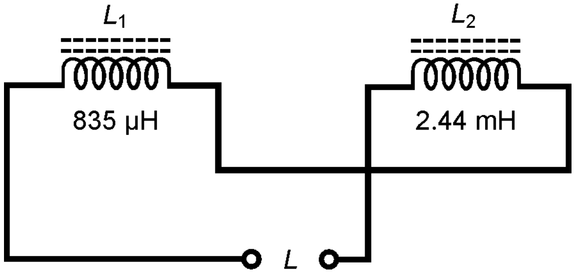


10-5 Illustration for Problem 10-5.

First, calculate M from k. According to the formula for this, given previously, $M = 0.500(50 \times 30)^{1/2} = 19.4$ μH. Then figure the total inductance. It is equal to $L = L_1 + L_2 + 2M = 30 + 50 + 38.8 = 118.8$ μH, rounded to 120 μH because only two significant digits are justified.

Problem 10-6

Imagine two coils with inductances of $L_1 = 835$ μH and $L_2 = 2.44$ mH. Suppose they are connected in series so that their coefficient of coupling is 0.922, acting so that the coils oppose each other, as shown in Fig. 10-6. What is the net inductance of the pair?



10-6 Illustration for Problem 10-6.

First, calculate M from k. The coil inductances are specified in different units. Let's use microhenrys for our calculations, so $L_2 = 2440$ µH. Then $M = 0.922(835 \times 2440)^{1/2} = 1316$ µH. Then figure the total inductance. It is $L = L_1 + L_2 - 2M = 835 + 2440 - 2632 = 643$ µH.

Air-Core Coils

The simplest inductors (besides plain, straight lengths of wire) are coils. A coil can be wound on a hollow cylinder of plastic or other nonferromagnetic material, forming an *air-core coil.* In practice, the maximum attainable inductance for such coils is about 1 mH.

Air-core coils are used mostly in radio-frequency transmitters, receivers, and antenna networks. In general, the higher the frequency of ac, the less inductance is needed to produce significant effects. Air-core coils can be made to have almost unlimited current-carrying capacity, simply by using heavy-gauge wire and making the radius of the coil large. Air does not dissipate much energy in the form of heat. It's efficient, even though it has low permeability.

Ferromagnetic Cores

Ferromagnetic substances can be crushed into dust and then bound into various shapes, providing core materials that greatly increase the inductance of a coil having a given number of turns. Depending on the mixture used, the increase in flux density can range from a factor of a few times, up through many thousands of times. A small coil can thus be made to have a large inductance. There are two main types of ferromagnetic material in common use as coil cores. These substances are known as *powdered iron* and *ferrite.*

Advantages and Limitations

Powdered-iron cores are common at high and very high radio frequencies. Ferrite is a special form of powdered iron that has exceptionally high permeability, causing a great concentration of magnetic flux lines within the coil. Ferrite is used at audio frequencies, as well as at low, medium, and high radio frequencies. Coils using these materials can be made much smaller, physically, than can air-core coils having the same inductance.

The main trouble with ferromagnetic cores is that, if the coil carries more than a certain amount of current, the core will *saturate.* This means that the ferromagnetic material is holding as much flux as it possibly can. When a core becomes saturated, any further increase in coil current will not produce a corresponding increase in the magnetic flux in the core. The result is that the inductance changes, decreasing with coil currents that are more than the critical value. In extreme cases, ferromagnetic cores can also waste considerable power as heat. This makes a coil *lossy.*

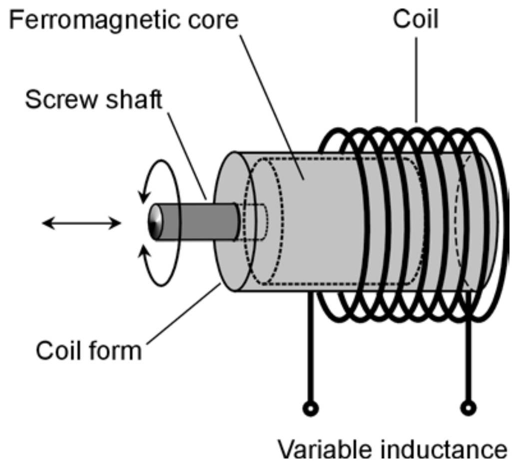
Permeability Tuning

Solenoidal coils can be made to have variable inductance by sliding ferromagnetic cores in and out of them. The frequency of a radio circuit can be adjusted in this way, as you'll learn later in this book.

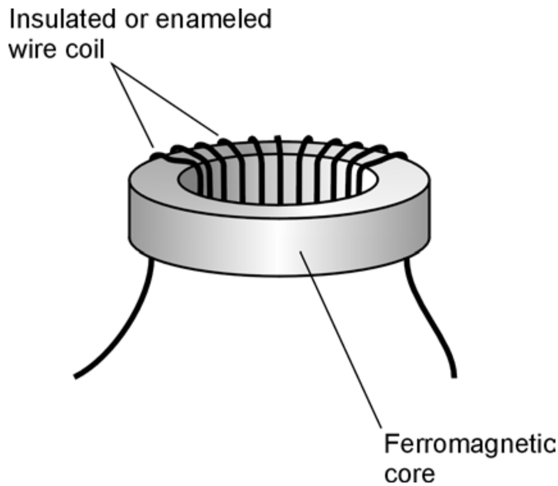
Because moving the core in and out of a coil changes the effective permeability within the coil, this method of tuning is called *permeability tuning.* The in/out motion can be precisely controlled by attaching the core to a screw shaft, and anchoring a nut at one end of the coil (Fig. 10-7). As the screw shaft is rotated clockwise, the core enters the coil, and the inductance increases. As the screw shaft is rotated counterclockwise, the core moves out of the coil, and the inductance decreases.

Toroids

Inductor coils do not have to be wound on cylindrical forms, or on cylindrical ferromagnetic cores. There's another coil geometry, called the *toroid.* It gets its name from the shape of the ferromagnetic core. The coil is wound over a core having this shape (Fig. 10-8), which resembles a donut or bagel.



10-7 Permeability tuning can be accomplished by moving a ferromagnetic core in and out of a solenoidal coil.



10-8 A toroidal coil is wound on a donut-shaped ferromagnetic core.

There are several advantages to toroidal coils over solenoidal, or cylindrical, ones. First, fewer turns of wire are needed to get a certain inductance with a toroid compared to a solenoid. Second, a toroid can be physically smaller for a given inductance and current-carrying capacity. Third, practically all the flux is contained within the core material. This reduces unwanted mutual inductances with components near the toroid.

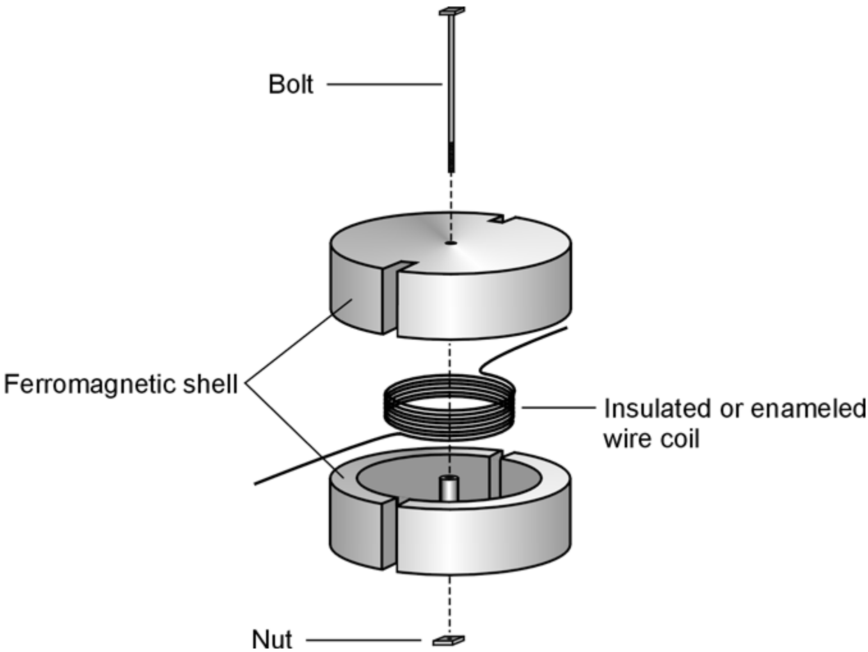
Toroidal coils have limitations, too. It is more difficult to permeability-tune a toroidal coil than it is to tune a solenoidal one. Toroidal coils are harder to wind than solenoidal ones. Sometimes, mutual inductance between or among physically separate coils is actually desired; with a toroid, the coils have to be wound on the same form for this to be possible.

Pot Cores

There is another way to confine the magnetic flux in a coil so that unwanted mutual inductance does not occur: wrap ferromagnetic core material around a coil (Fig. 10-9). A wraparound core of this sort is known as a *pot core.*

A typical pot core comes in two halves, inside one of which the coil is wound. Then the parts are assembled and held together by a bolt and nut. The entire assembly looks like a miniature oil tank. The wires come out of the core through small holes or slots.

Pot cores have the same advantages as toroids. The core tends to prevent the magnetic flux from extending outside the physical assembly. Inductance is greatly increased compared to solenoidal windings having a comparable number of turns. In fact, pot cores are even better than toroids if the main objective is to get a large inductance in a small space. The main disadvantage of a pot core is that tuning, or adjustment of the inductance, is all but impossible. The only way to do it is by switching in different numbers of turns, using taps at various points on the coil.



10-9 Exploded view of a pot core. The coil winding is inside the ferromagnetic shell.

Filter Chokes

The largest values of inductance that can be obtained in practice are on the order of several henrys. The primary use of a coil this large is to smooth out the pulsations in direct current that result when ac is *rectified* in a power supply. This type of coil is known as a *filter choke.* You'll learn more about power supplies later in this book.

Inductors at AF

Inductors for audio frequency (AF) applications range in value from a few millihenrys up to about 1 H. They are almost always toroidally wound, or are wound in a pot core, or comprise part of an audio transformer. Ferromagnetic cores are the rule.

Inductors can be used in conjunction with moderately large values of capacitance in order to obtain *AF-tuned circuits.* However, in recent years, audio tuning has been largely taken over by active components, particularly *integrated circuits.*

Inductors at RF

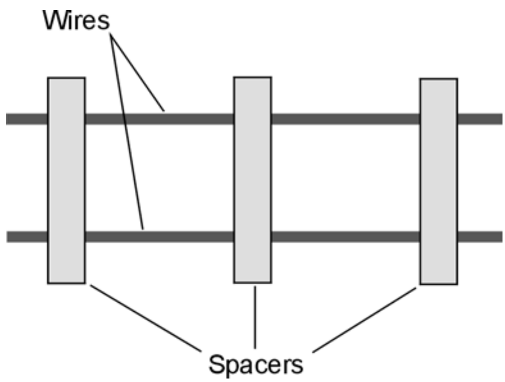
The radio frequency (RF) spectrum ranges from a few kilohertz to well above 100 GHz. At the low end of this range, inductors are similar to those at AF. As the frequency increases, cores having lower permeability are used. Toroids are common up through about 30 MHz. Above that frequency, air-core coils are more often used.

In RF applications, coils are routinely connected in series or in parallel with capacitors to obtain tuned circuits. Other arrangements yield various characteristics of *attenuation versus frequency*, serving to let signals at some frequencies pass through, while rejecting signals at other frequencies. You'll learn more about this in the discussion about *resonance* in Chap. 17.

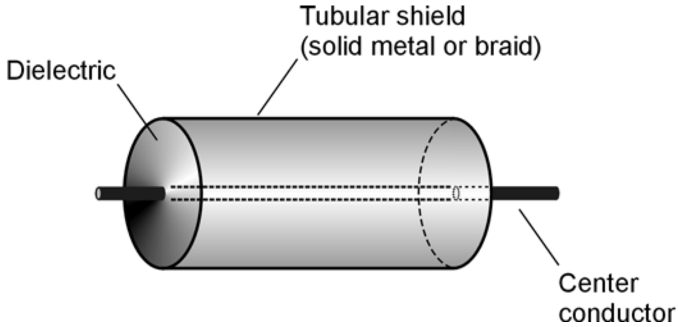
Transmission-Line Inductors

At frequencies about 100 MHz, another type of inductor becomes practical. This is the type formed by a length of *transmission line.* A transmission line is generally used to get energy from one place to another. In radio communications, transmission lines get energy from a transmitter to an antenna, and from an antenna to a receiver.

Most transmission lines are found in either of two geometries, the *parallel-wire* type or the *coaxial* type. A parallel-wire transmission line consists of two wires running alongside each other with constant spacing (Fig. 10-10). The spacing is maintained by polyethylene rods molded at regular in-



10-10 Parallel-wire transmission line. The spacers are made of sturdy insulating material.



10-11 Coaxial transmission line. The dielectric material keeps the center conductor along the axis of the tubular shield.

tervals to the wires, or by a solid web of polyethylene. The substance separating the wires is called the *dielectric* of the transmission line. A coaxial transmission line has a wire conductor surrounded by a tubular braid or pipe (Fig. 10-11). The wire is kept at the center of this tubular *shield* by means of polyethylene beads, or more often, by solid or foamed polyethylene, all along the length of the line.

Line Inductance

Short lengths of any type of transmission line behave as inductors, as long as the line length is less than 90° (1⁄4 of a wavelength). At 100 MHz, 90° in free space is 75 cm, or a little more than 2 ft. In general, if f is the frequency in megahertz, then 1⁄4 wavelength in free space, expressed in centimeters (s_{cm}), is given by this formula:

$$s_{cm} = 7500/f$$

The length of a quarter-wavelength section of transmission line is shortened from the free-space quarter wavelength by the effects of the dielectric. In practice, 1⁄4 wavelength along the line can be anywhere from about 0.66 (or 66 percent) of the free-space length for coaxial lines with solid polyethylene dielectric to about 0.95 (or 95 percent) of the free-space length for parallel-wire line with spacers molded at intervals of several centimeters. The factor by which the wavelength is shortened is called the *velocity factor* of the line.

The shortening of the wavelength in a transmission line, compared with the wavelength in free space, is a result of a slowing down of the speed with which the radio signals move in the line compared with their speed in space (the speed of light). If the velocity factor of a line is given by v, then the preceding formula for the length of a quarter-wave line, in centimeters, becomes:

$$s_{cm} = 7500v/f$$

Very short lengths of line—a few electrical degrees—produce small values of inductance. As the length approaches 1⁄4 wavelength, the inductance increases.

Transmission line inductors behave differently than coils in one important way: the inductance of a coil, particularly an air-core coil, is independent of the frequency. But the inductance of a transmission-line section changes as the frequency changes. At first, the inductance becomes larger as the frequency increases. At a certain limiting frequency, the inductance becomes theoretically infinite. Above that frequency, the line becomes *capacitive* rather than inductive. You'll learn about capacitance in the next chapter.

Unwanted Inductances

Any length of wire has some inductance. As with a transmission line, the inductance of a wire increases as the frequency increases. Wire inductance is more significant at RF than at AF.

In some cases, especially in radio communications equipment, the inductance of, and among, wires can become a major problem. Circuits can oscillate when they should not. A receiver might respond to signals that it's not designed to intercept. A transmitter can send out signals on unauthorized and unintended frequencies. The frequency response of any circuit can be altered, degrading the performance of the equipment. Sometimes the effects of this *stray inductance* are so small that they are not important; this might be the case in a stereo hi-fi set located at a distance from other electronic equipment. But in some situations, stray inductance can cause serious equipment malfunctions.

A good way to minimize stray inductance is to use coaxial cables between and among sensitive circuits or components. The shield of the cable is connected to the *common ground* of the apparatus. In some cases, enclosing individual circuits in metal boxes can prevent stray inductance from causing feedback and other problems.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. An inductor works by
   (a) charging a piece of wire.
   (b) storing energy as a magnetic field.
   (c) choking off dc.
   (d) introducing resistance into a circuit.

2. Which of the following does *not* affect the inductance of an air-core coil, if all other factors are held constant?
   (a) The frequency
   (b) The number of turns
   (c) The diameter of the coil
   (d) The length of the coil

3. In a small inductance
   (a) energy is stored and released slowly.
   (b) the current flow is always large.
   (c) the current flow is always small.
   (d) energy is stored and released quickly.

4. A ferromagnetic core is placed in an inductor mainly to
   (a) increase the current carrying capacity.
   (b) increase the inductance.
   (c) limit the current.
   (d) reduce the inductance.

5. Inductors in series, assuming there is no mutual inductance, combine
   (a) like resistors in parallel.
   (b) like resistors in series.
   (c) like batteries in series with opposite polarities.
   (d) in a way unlike any other type of component.

6. Suppose two inductors are connected in series, without mutual inductance. Their values are 33 mH and 55 mH. What is the net inductance of the combination?
   (a) 1.8 H
   (b) 22 mH
   (c) 88 mH
   (d) 21 mH

7. If the same two inductors (33 mH and 55 mH) are connected in parallel without mutual inductance, the combination will have a value of
   (a) 1.8 H.
   (b) 22 mH.
   (c) 88 mH.
   (d) 21 mH.

8. Suppose three inductors are connected in series without mutual inductance. Their values are 4.00 nH, 140 μH, and 5.07 H. For practical purposes, the net inductance will be very close to
   (a) 4.00 nH.
   (b) 140 μH.
   (c) 5.07 H.
   (d) none of the above.

9. Suppose the three inductors mentioned above are connected in parallel without mutual inductance. The net inductance will be close to
   (a) 4.00 nH.
   (b) 140 μH.
   (c) 5.07 H.
   (d) none of the above.

10. Suppose two inductors, each of 100 μH, are connected in series, and the coefficient of coupling is 0.40. The net inductance, if the coil fields reinforce each other, is
    (a) 50.0 μH.
    (b) 120 μH.
    (c) 200 μH.
    (d) 280 μH.

11. If the coil fields oppose in the foregoing series-connected arrangement, assuming the coefficient of coupling does not change, the net inductance is
    (a) 50.0 μH.
    (b) 120 μH.
    (c) 200 μH.
    (d) 280 μH.

12. Suppose two inductors, having values of 44.0 mH and 88.0 mH, are connected in series with a coefficient of coupling equal to 1.0 (the maximum possible mutual inductance). If their fields reinforce, the net inductance is approximately
    (a) 7.55 mH.
    (b) 132 mH.
    (c) 194 mH.
    (d) 256 mH.

13. If the fields in the previous situation oppose, assuming the coefficient of coupling does not change, the net inductance will be approximately
    (a) 7.55 mH.
    (b) 132 mH.
    (c) 194 mH.
    (d) 256 mH.

14. With permeability tuning, moving the core further into a solenoidal coil
    (a) increases the inductance.
    (b) reduces the inductance.
    (c) has no effect on the inductance, but increases the current-carrying capacity of the coil.
    (d) raises the frequency.

15. A significant advantage, in some situations, of a toroidal coil over a solenoid is the fact that
    (a) the toroid is easier to wind.
    (b) the solenoid cannot carry as much current.
    (c) the toroid is easier to tune.
    (d) the magnetic flux in a toroid is practically all within the core.

16. A major feature of a pot core inductor is
    (a) high current capacity.
    (b) large inductance in small volume.
    (c) excellent efficiency at very high frequencies.
    (d) ease of inductance adjustment.

17. As an inductor core material, air
    (a) has excellent efficiency.
    (b) has high permeability.
    (c) allows large inductance to exist in a small volume.
    (d) has permeability that can vary over a wide range.

18. At a frequency of 400 Hz, which is in the AF range, the most likely form for an inductor would be
    (a) air-core.
    (b) solenoidal.
    (c) toroidal.
    (d) transmission-line.

19. At a frequency of 95.7 MHz, which is in the frequency-modulation (FM) broadcast band and is considered part of the very high frequency (VHF) radio spectrum, a good form for an inductor would be
    (a) air-core.
    (b) pot core.
    (c) either (a) or (b).
    (d) neither (a) nor (b).

20. A transmission-line inductor made from coaxial cable having velocity factor of 0.66 and working at 450 MHz, which is in the ultrahigh frequency (UHF) radio spectrum, should, in order to measure less than ¼ electrical wavelength, be cut shorter than
    (a) 16.7 m.
    (b) 11 m.
    (c) 16.7 cm.
    (d) 11 cm.

11
CHAPTER

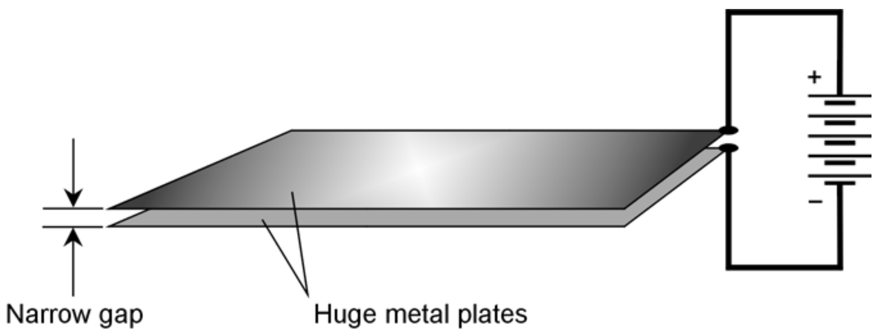
Capacitance

ELECTRICAL COMPONENTS CAN OPPOSE THE FLOW OF AC IN THREE WAYS, TWO OF WHICH YOU'VE learned about. *Resistance* slows the flow of ac or dc charge carriers (usually electrons) by brute force. *Inductance* impedes the flow of ac charge carriers by temporarily storing the energy as a magnetic field. *Capacitance*, about which you'll learn in this chapter, impedes the flow of ac charge carriers by temporarily storing the energy as an *electric field.*

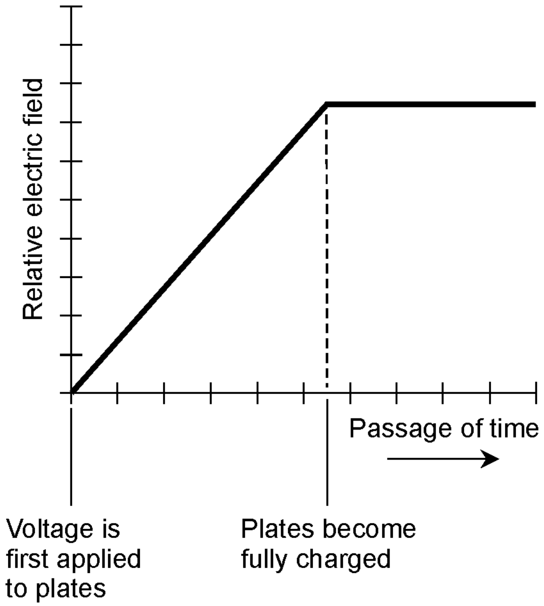
The Property of Capacitance

Imagine two huge, flat sheets of metal that are excellent electrical conductors. Suppose they are each the size of the state of Nebraska, and are placed one over the other, separated by only 1 foot of space. If these two sheets of metal are connected to the terminals of a battery, as shown in Fig. 11-1, they will become charged electrically, one positively and the other negatively.

If the plates were small, they would both become charged almost instantly, attaining a relative voltage equal to the voltage of the battery. But because the plates are gigantic, it will take a little time for the negative plate to reach full negative potential, and an equal time for the other plate to reach full positive potential. Eventually, the voltage between the two plates will equal the battery voltage,



11-1 A hypothetical gigantic capacitor.



11-2 Relative electric field intensity between metal plates connected to a voltage source, as a function of time.

and an electric field will exist in the space between the plates. This electric field will be small at first, because the plates don't charge up right away. But the charge will increase over a period of time, depending on how large the plates are, and also depending on how far apart they are. Figure 11-2 is a relative graph showing the intensity of the electric field between the plates as a function of time, elapsed from the instant the plates are connected to the battery terminals.

Energy will be stored in this electric field. The ability of the plates, and of the space between them, to store this energy is the property of *capacitance.* As a quantity or variable, capacitance is denoted by the uppercase italic letter *C.*

Practical Capacitors

It's out of the question to make a capacitor of the preceding dimensions. But two sheets, or strips, of foil can be placed one on top of the other, separated by a thin, nonconducting sheet such as paper, and then the whole assembly can be rolled up to get a large effective surface area. When this is done, the electric flux becomes great enough so that the device exhibits significant capacitance. Alternatively, two sets of several plates each can be meshed together with air in between them, and the resulting capacitance is significant at high ac frequencies.

In a capacitor, the electric flux concentration is multiplied when a *dielectric* of a certain type is placed between the plates. This increases the effective surface area of the plates, so that a physically small component can be made to have a large capacitance. The voltage that a capacitor can handle depends on the thickness of the metal sheets or strips, on the spacing between them, and on the type of dielectric used.

In general, capacitance is directly proportional to the surface area of the conducting plates or sheets. Capacitance is *inversely proportional* to the separation between conducting sheets. In other words, the closer the sheets are to each other, the greater the capacitance. The capacitance also depends on the *dielectric constant* of the material between the plates. A vacuum has a dielectric constant of 1; some substances have dielectric constants that multiply the effective capacitance many times.

The Unit of Capacitance

When a battery is connected between the plates of a capacitor, the potential difference between the plates builds up at a rate that depends on the capacitance. The greater the capacitance, the slower the rate of change of voltage in the plates. The unit of capacitance is an expression of the ratio between the current that flows and the rate of voltage change between the plates as the plates become charged. A capacitance of 1 *farad* (1 F) represents a current flow of 1 A while there is a voltage increase of 1 V/s. A capacitance of 1 F also results in 1 V of potential difference for an electric charge of 1 C.

The farad is a huge unit of capacitance. You'll almost never see a capacitor with a value of 1 F. Commonly employed units of capacitance are the *microfarad* (μF) and the *picofarad* (pF). A capacitance of 1 μF represents 0.000001 (10^{-6}) F, and 1 pF is a millionth of a microfarad, or 0.000000000001 (10^{-12}) F.

Physically small components can be made to have fairly large capacitance values. Conversely, some capacitors with small values take up large physical volumes. The physical size of a capacitor, if all other factors are held constant, is proportional to the voltage that it can handle. The higher the rated voltage, the bigger the component.

Capacitors in Series

With capacitors, there is rarely any mutual interaction. This makes capacitors easier to work with than inductors. We don't have to worry about *mutual capacitance* very often, the way we have to be concerned about mutual inductance when working with wire coils.

Capacitors in series add together like resistors or inductors in parallel. Suppose you have several capacitors with values C_1, C_2, C_3, . . . , C_n connected in series. You can find the reciprocal of the total capacitance, $1/C$, using the following formula:

$$1/C = 1/C_1 + 1/C_2 + 1/C_3 + \ldots + 1/C_n$$

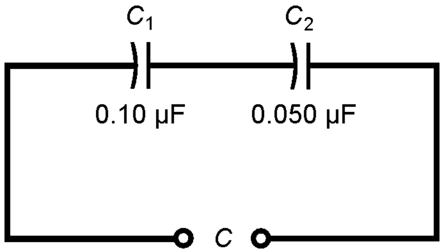
The net capacitance of the series combination, C, is found by taking the reciprocal of the number you get for $1/C$.

If two or more capacitors are connected in series, and one of them has a value that is tiny compared with the values of all the others, the net capacitance is roughly equal to the smallest capacitance.

Problem 11-1

Suppose two capacitors, with values of $C_1 = 0.10$ μF and $C_2 = 0.050$ μF, are connected in series (Fig. 11-3). What is the net capacitance?

Using the preceding formula, first find the reciprocals of the values. They are $1/C_1 = 10$ and $1/C_2 = 20$. Then $1/C = 10 + 20 = 30$, and $C = 1/30 = 0.033$ μF. Note that we can work with reciprocal capacitances in this calculation only because the values of the components are specified in the same units.



11-3 Capacitors in series. Illustration for Problem 11-1.

Problem 11-2

Suppose two capacitors with values of 0.0010 μF and 100 pF are connected in series. What is the net capacitance?

In this case, you must convert to the same size units before doing any calculations. A value of 100 pF represents 0.000100 μF. Thus, $C_1 = 0.0010$ μF and $C_2 = 0.000100$ μF. The reciprocals are $1/C_1 = 1000$ and $1/C_2 = 10{,}000$. Therefore, $1/C = 1000 + 10{,}000 = 11{,}000$, so $C = 1/11{,}000 = 0.000091$ μF. (You might rather say it's 91 pF.)

Problem 11-3

Suppose five capacitors, each of 100 pF, are in series. What is the total capacitance?

If there are n capacitors in series, all of the same value so that $C_1 = C_2 = C_3 = \ldots = C_n$, the net capacitance C is equal to $1/n$ of the capacitance of any of the components alone. Because there are five 100-pF capacitors here, the total is $C = 100/5 = 20.0$ pF.

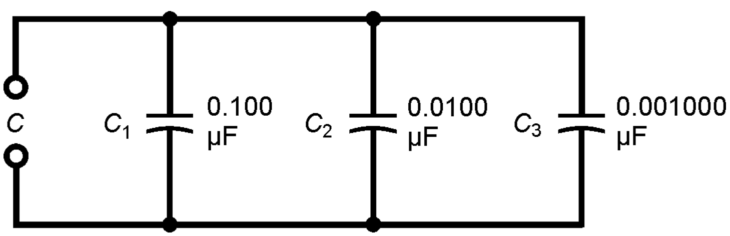
Capacitors in Parallel

Capacitances in parallel add like resistances in series. The total capacitance is the sum of the individual component values. If two or more capacitors are connected in parallel, and one of the capacitances is far larger than any of the others, the total capacitance can be taken as approximately the value of the biggest one.

Problem 11-4

Suppose three capacitors are in parallel, having values of $C_1 = 0.100$ μF, $C_2 = 0.0100$ μF, and $C_3 = 0.001000$ μF, as shown in Fig. 11-4. What is the total capacitance?

Add them up: $C = 0.100 + 0.0100 + 0.001000 = 0.111000$. Because two of the values are given to only three significant figures, the final answer should be stated as $C = 0.111$ μF.



11-4 Capacitors in parallel. Illustration for Problem 11-4.

Problem 11-5

Suppose two capacitors are in parallel, one with a value of 100 μF and one with a value of 100 pF. What is the net capacitance?

In this case, you can say right away that the net capacitance is 100 μF for practical purposes. The 100-pF capacitor has a value that is only one-millionth of the capacitance of the 100-μF component. The smaller capacitance contributes essentially nothing to the net capacitance of this combination.

Fixed Capacitors

A *fixed capacitor* has a value that cannot be adjusted, and that (ideally) does not vary when environmental or circuit conditions change. Here are some of the characteristics, and common types, of fixed capacitors.

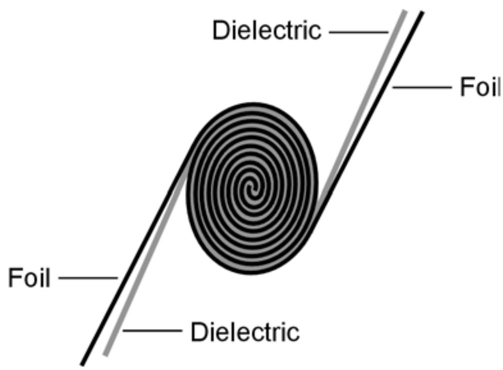
Dielectric Materials

Just as certain solids can be placed within a coil to increase the inductance, materials exist that can be sandwiched in between the plates of a capacitor to increase the capacitance. The substance between the plates is called the *dielectric* of the capacitor. Air is an efficient dielectric; it has almost no loss. But it is difficult to get very much capacitance using air as the dielectric. Some kind of solid material is usually employed as the dielectric for most fixed capacitors.

Dielectric materials accommodate electric *fields* well, but they are poor conductors of electric *currents.* In fact, dielectric materials are known as good insulators. Solid dielectrics increase the capacitance for a given surface area and spacing of the plates. Solid dielectrics also allow the plates to be rolled up, squashed, and placed very close together (Fig. 11-5). This geometry acts to maximize the capacitance per unit volume.

Paper Capacitors

In the early days of electronics, capacitors were commonly made by placing paper, soaked with mineral oil, between two strips of foil, rolling the assembly up, attaching wire leads to the two pieces of foil, and enclosing the rolled-up foil and paper in an airtight cylindrical case. *Paper capacitors* can still sometimes be found in older electronic equipment. They have values ranging from about 0.001 μF to 0.1 μF, and can handle low to moderate voltages, usually up to about 1000 V.



11-5 A cross-sectional drawing of a capacitor consisting of two foil sheets rolled up, and two sheets of dielectric material rolled up between them.

Mica Capacitors

Mica is a naturally occurring, solid, transparent mineral substance that flakes off in thin sheets. It makes an excellent dielectric for capacitors. *Mica capacitors* can be manufactured by alternately stacking metal sheets and layers of mica, or by applying silver ink to sheets of mica. The metal sheets are wired together into two meshed sets, forming the two terminals of the capacitor. This scheme is shown in Fig. 11-6.

Mica capacitors have low loss, and are therefore highly efficient, provided their voltage rating is not exceeded. Voltage ratings can be up to several thousand volts if thick sheets of mica are used. But mica capacitors are large physically in proportion to their capacitance. The main application for mica capacitors is in radio receivers and transmitters. Their capacitances are a little lower than those of paper capacitors, ranging from a few tens of picofarads up to about 0.05 μF.

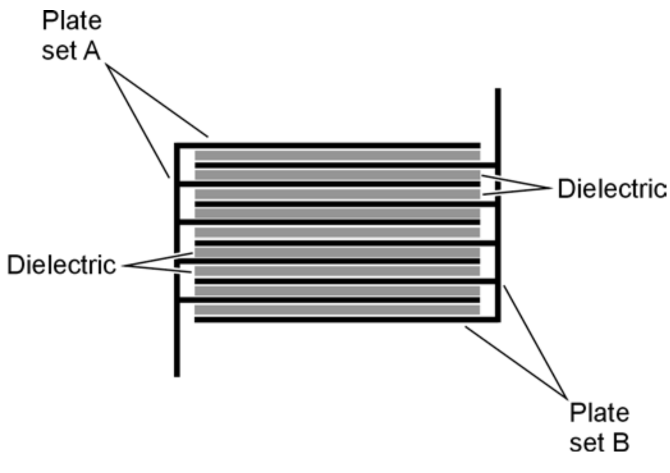
Ceramic Capacitors

Ceramic materials work well as dielectrics. Sheets of metal are stacked alternately with wafers of ceramic to make these capacitors. The meshing/layering geometry of Fig. 11-6 is used. Ceramic, like mica, has low loss and allows for high efficiency.

For small values of capacitance, only one layer of ceramic is needed, and two metal plates can be glued to the disk-shaped material, one on each side. This type of component is known as a *disk-ceramic* capacitor. Alternatively, a tube or cylinder of ceramic can be employed, and metal ink applied to the inside and outside of the tube. Such units are called *tubular* capacitors. Ceramic capacitors have values ranging from a few picofarads to about 0.5 μF. Their voltage ratings are comparable to those of paper capacitors.

Plastic-Film Capacitors

Plastics make good dielectrics for the manufacture of capacitors. *Polyethylene* and *polystyrene* are commonly used. The method of manufacture is similar to that for paper capacitors. Stacking methods can be used if the plastic is rigid. The geometries can vary, and these capacitors are therefore found in various shapes.



11-6 A cross-sectional drawing of a capacitor consisting of two meshed sets of several metal plates, separated by layers of dielectric material.

Capacitance values for plastic-film units range from about 50 pF to several tens of microfarads. Most often they are in the range of 0.001 μF to 10 μF. Plastic capacitors are employed at AF and RF, and at low to moderate voltages. The efficiency is good, although not as high as that for mica-dielectric or air-dielectric units.

Electrolytic Capacitors

All of the aforementioned types of capacitors provide relatively small values of capacitance. They are also *nonpolarized*, meaning that they can be hooked up in a circuit in either direction. An *electrolytic* capacitor provides greater capacitance than any of the preceding types, but it must be connected in the proper direction in a circuit to work right. An electrolytic capacitor is a *polarized* component.

Electrolytic capacitors are made by rolling up aluminum foil strips, separated by paper saturated with an *electrolyte* liquid. The electrolyte is a conducting solution. When dc flows through the component, the aluminum oxidizes because of the electrolyte. The oxide layer is nonconducting, and forms the dielectric for the capacitor. The layer is extremely thin, and this results in a high capacitance per unit volume. Electrolytic capacitors can have values up to thousands of microfarads, and some can handle thousands of volts. These capacitors are most often seen in AF circuits and in dc power supplies.

Tantalum Capacitors

Another type of electrolytic capacitor uses tantalum rather than aluminum. The tantalum can be foil, as is the aluminum in a conventional electrolytic capacitor. It can also take the form of a porous pellet, the irregular surface of which provides a large area in a small volume. An extremely thin oxide layer forms on the tantalum.

Tantalum capacitors have high reliability and excellent efficiency. They are often used in military applications because they almost never fail. They can be used in AF and digital circuits in place of aluminum electrolytics.

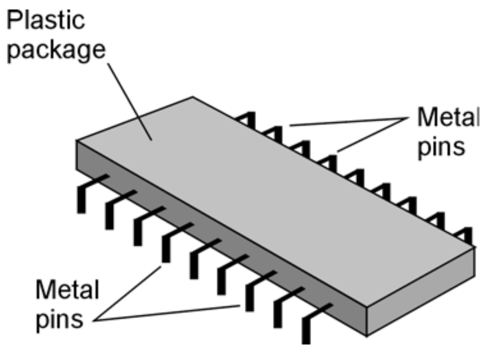
Semiconductor Capacitors

Later in this book, you'll learn about *semiconductors.* These materials have revolutionized electrical and electronic circuit design in the past several decades.

Semiconductor materials can be employed to make capacitors. A semiconductor *diode* conducts current in one direction, and refuses to conduct in the other direction. When a voltage source is connected across a diode so that it does not conduct, the diode acts as a capacitor. The capacitance varies depending on how much of this *reverse voltage* is applied to the diode. The greater the reverse voltage, the smaller the capacitance. This makes the diode act as a *variable capacitor.* Some diodes are especially manufactured to serve this function. Their capacitances fluctuate rapidly along with pulsating dc. They are called *varactor diodes* or simply *varactors.*

Capacitors can be formed in the semiconductor materials of an integrated circuit (also called an *IC* or *chip*) in much the same way. Sometimes, IC diodes are fabricated to serve as varactors. Another way to make a capacitor in an IC is to sandwich an oxide layer into the semiconductor material, between two layers that conduct well. Most ICs look like little boxes with protruding metal prongs (Fig. 11-7). The prongs provide the electrical connections to external circuits and systems.

Semiconductor capacitors usually have small values of capacitance. They are physically tiny, and can handle only low voltages. The advantages are miniaturization and an ability, in the case of the varactor, to change in value at a rapid rate.



11-7 A typical integrated-circuit package is a tiny plastic box with protruding metal pins.

Variable Capacitors

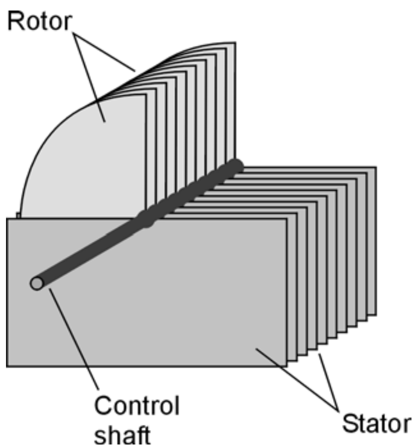
The capacitance of a component can be varied at will by adjusting the mutual surface area between the plates, or by changing the spacing between the plates. The two most common types of variable capacitors (besides varactors) are the *air variable* and the *trimmer.* You will also sometimes encounter *coaxial capacitors.*

Air Variables

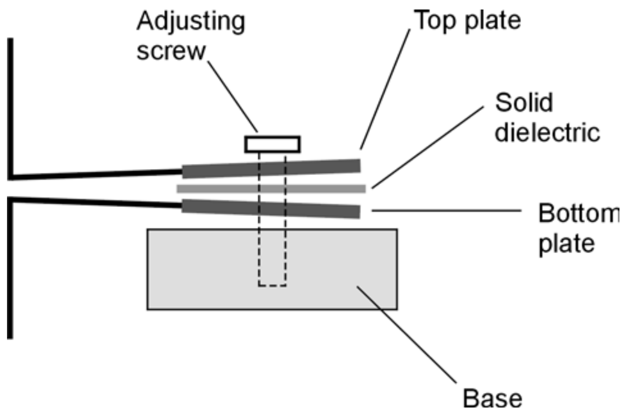
By connecting two sets of metal plates so that they mesh, and by affixing one set to a rotatable shaft, a variable capacitor is made. The rotatable set of plates is called the *rotor*, and the fixed set is called the *stator.* This is the type of component you might have seen in older radio receivers, used to tune the frequency. Such capacitors are still used in transmitter output tuning networks. Figure 11-8 is a functional drawing of an air-variable capacitor.

Air variables have maximum capacitance that depends on the number of plates in each set, and also on the spacing between the plates. Common maximum values are 50 to 500 pF; minimum values are a few picofarads. The voltage-handling capability depends on the spacing between the plates. Some air variables can handle many kilovolts.

Air variables are used primarily in RF applications. They are highly efficient, and are nonpolarized, although the rotor is usually connected to common ground (the chassis or circuit board).



11-8 A simplified drawing of an air-variable capacitor.



11-9 A cross-sectional drawing of a trimmer capacitor.

Trimmer Capacitors

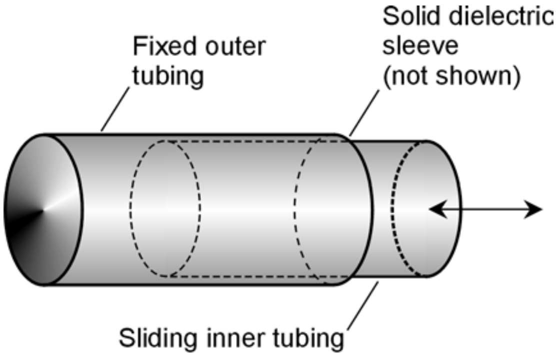
When it is not necessary to change the value of a capacitor very often, a trimmer can be used. It consists of two plates, mounted on a ceramic base and separated by a sheet of plastic, mica, or some other solid dielectric. The plates are flexible, and can be squashed together more or less by means of a screw (Fig. 11-9). Sometimes two sets of several plates are interleaved to increase the capacitance.

Trimmers can be connected in parallel with an air variable, so that the range of the air variable can be adjusted. Some air-variable capacitors have trimmers built in. Typical maximum values for trimmers range from a few picofarads up to about 200 pF. They handle low to moderate voltages, are highly efficient, and are nonpolarized.

Coaxial Capacitors

You recall from the previous chapter that sections of transmission lines can work as inductors. They can act as capacitors, too. If a section of transmission line is less than $\frac{1}{4}$ wavelength long, and is left open at the far end (rather than shorted out), it behaves as a capacitor. The capacitance increases with length.

The most common transmission-line capacitor uses two telescoping sections of metal tubing. This is called a *coaxial capacitor.* It works because there is a certain effective surface area between the inner and the outer tubing sections. A sleeve of plastic dielectric is placed between the sections of tubing, as shown in Fig. 11-10. This allows the capacitance to be adjusted by sliding the inner section in or out of the outer section.



11-10 A simplified drawing of a coaxial variable capacitor.

Coaxial capacitors are used in RF applications, particularly in antenna systems. Their values are generally from a few picofarads up to about 100 pF.

Capacitor Specifications

When you are looking for a capacitor for a particular application, it's important to find a component that has the right specifications for the job. Here are two of the most important specifications to watch for.

Tolerance

Capacitors are rated according to how nearly their values can be expected to match the rated capacitance. The most common tolerance is $\pm 10\%$; some capacitors are rated at $\pm 5\%$ or even at $\pm 1\%$.

The lower (or *tighter*) the tolerance number, the more closely you can expect the actual component value to match the rated value. For example, a $\pm 10\%$ capacitor rated at 100 pF can range from 90 to 110 pF. But if the tolerance is $\pm 1\%$, the manufacturer guarantees that the capacitance will be between 99 and 101 pF.

Problem 11-6

A capacitor is rated at 0.10 μF $\pm 10\%$. What is its guaranteed range of capacitance?

First, multiply 0.10 by 10 percent to get the plus-or-minus variation. This is $0.10 \times 0.10 = 0.010$ μF. Then add and subtract this from the rated value to get the maximum and minimum possible capacitances. The result is a range of 0.09 to 0.11 μF.

Temperature Coefficient

Some capacitors increase in value as the temperature increases. These components have a *positive temperature coefficient.* Some capacitors decrease in value as the temperature rises; these have a *negative temperature coefficient.* Some capacitors are manufactured so that their values remain constant over a certain temperature range. Within this span of temperatures, such capacitors have *zero temperature coefficient.*

The temperature coefficient is specified in percent per degree Celsius (%/°C). Sometimes, a capacitor with a negative temperature coefficient can be connected in series or parallel with a capacitor having a positive temperature coefficient, and the two opposite effects cancel out over a range of temperatures. In other instances, a capacitor with a positive or negative temperature coefficient can be used to cancel out the effect of temperature on other components in a circuit, such as inductors and resistors.

Interelectrode Capacitance

Any two pieces of conducting material, when they are brought near each other, can act as a capacitor. Often, this *interelectrode capacitance* is so small that it can be neglected. It rarely amounts to more than a few picofarads. In utility circuits and at AF, interelectrode capacitance is not usually significant. But it can cause problems at RF. The chances for trouble increase as the frequency increases. The most common phenomena are *feedback*, and/or a change in the frequency characteristics of a circuit.

Interelectrode capacitance can be minimized by keeping wire leads as short as possible, by using shielded cables, and by enclosing sensitive circuits in metal housings.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. Capacitance acts to store electrical energy as
   (a) current.
   (b) voltage.
   (c) a magnetic field.
   (d) an electric field.
2. As capacitor plate area increases, all other things being equal,
   (a) the capacitance increases.
   (b) the capacitance decreases.
   (c) the capacitance does not change.
   (d) the current-handling ability decreases.
3. As the spacing between plates in a capacitor is made smaller, all other things being equal,
   (a) the capacitance increases.
   (b) the capacitance decreases.
   (c) the capacitance does not change.
   (d) the resistance increases.
4. A material with a high dielectric constant
   (a) acts to increase capacitance per unit volume.
   (b) acts to decrease capacitance per unit volume.
   (c) has no effect on capacitance.
   (d) causes a capacitor to become polarized.
5. A capacitance of 100 pF is the same as which of the following?
   (a) 0.01 μF
   (b) 0.001 μF
   (c) 0.0001 μF
   (d) 0.00001 μF
6. A capacitance of 0.033 μF is the same as which of the following?
   (a) 33 pF
   (b) 330 pF
   (c) 3300 pF
   (d) 33,000 pF
7. If five 0.050-μF capacitors are connected in parallel, what is the net capacitance of the combination?
   (a) 0.010 μF
   (b) 0.25 μF

(c) 0.50 μF
(d) 0.025 μF

8. If five 0.050-μF capacitors are connected in series, what is the net capacitance of the combination?
(a) 0.010 μF
(b) 0.25 μF
(c) 0.50 μF
(d) 0.025 μF

9. Suppose that two capacitors are connected in series, and their values are 47 pF and 33 pF. What is the net capacitance of this combination?
(a) 80 pF
(b) 47 pF
(c) 33 pF
(d) 19 pF

10. Suppose that two capacitors are in parallel. Their values are 47.0 pF and 470 μF. What is the net capacitance of this combination?
(a) 47.0 pF
(b) 517 pF
(c) 517 μF
(d) 470 μF

11. Suppose that three capacitors are in parallel. Their values are 0.0200 μF, 0.0500 μF, and 0.10000 μF. What is the net capacitance of this combination?
(a) 0.0125 μF
(b) 0.1700 μF
(c) 0.1000 μF
(d) 0.1250 μF

12. The main advantage of air as a dielectric material for capacitors is the fact that it
(a) has a high dielectric constant.
(b) is not physically dense.
(c) has low loss.
(d) allows for large capacitance in a small volume.

13. Which of the following is *not* a characteristic of mica capacitors?
(a) Excellent efficiency
(b) Small size, even for large values of capacitance
(c) High voltage-handling capacity
(d) Low loss

14. Which of the following capacitance values is most typical of a disk-ceramic capacitor?
    (a) 100 pF
    (b) 33 μF
    (c) 470 μF
    (d) 10,000 μF

15. Which of the following capacitance values is most typical of a paper capacitor?
    (a) 0.001 pF
    (b) 0.01 μF
    (c) 100 μF
    (d) 3300 μF

16. Which of the following capacitance ranges is most typical of an air-variable capacitor?
    (a) 0.01 μF to 1 μF
    (b) 1 μF to 100 μF
    (c) 1 pF to 100 pF
    (d) 0.001 pF to 0.1 pF

17. Which of the following types of capacitors is polarized?
    (a) Paper
    (b) Mica
    (c) Interelectrode
    (d) Electrolytic

18. If a capacitor has a negative temperature coefficient, then
    (a) its capacitance decreases as the temperature rises.
    (b) its capacitance increases as the temperature rises.
    (c) its capacitance does not change with temperature.
    (d) it will not work if the temperature is below freezing.

19. Suppose that a capacitor is rated at 33 pF $\pm$10%. Which of the following actual capacitance values is outside the acceptable range?
    (a) 30 pF
    (b) 37 pF
    (c) 35 pF
    (d) 31 pF

20. Suppose that a capacitor, rated at 330 pF, shows an actual value of 317 pF. By how many percent does its actual capacitance differ from its rated capacitance?
    (a) −0.039%
    (b) −3.9%
    (c) −0.041%
    (d) −4.1%

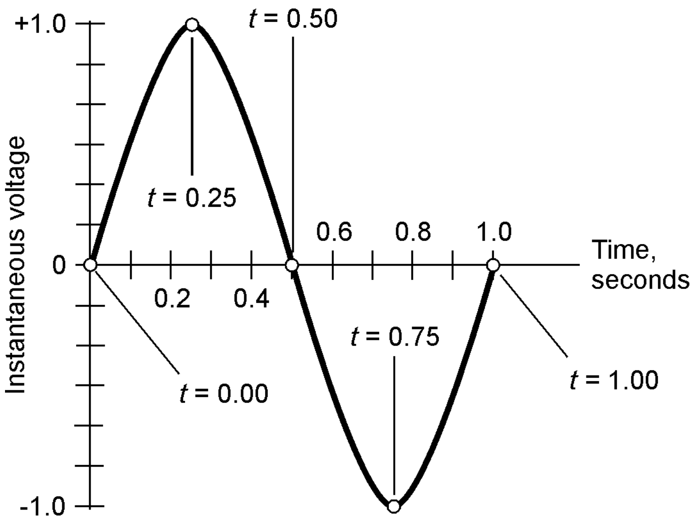
12
CHAPTER

Phase

IN ALTERNATING CURRENT, EACH 360° CYCLE IS EXACTLY THE SAME AS EVERY OTHER. IN EVERY CYCLE, the waveform of the previous cycle is repeated. In this chapter, you'll learn about the most common type of ac waveform: the *sine wave.*

Instantaneous Values

An ac sine wave has a characteristic shape, as shown in Fig. 12-1. This is the way the graph of the function $y = \sin x$ looks on an (x,y) coordinate plane. (The abbreviation *sin* stands for *sine* in trigonometry.) Suppose that the peak voltage is ± 1 V, as shown. Further imagine that the period is 1 s, so the frequency is 1 Hz. Let the wave begin at time $t = 0$. Then each cycle begins every time the value of t is a whole number. At every such instant, the voltage is zero and *positive-going.*



12-1 A sine wave with a period of 1 second. It thus has a frequency of 1 Hz.

If you freeze time at, say, $t = 446.00$, the voltage is zero. Looking at the diagram, you can see that the voltage will also be zero every so-many-and-a-*half* seconds, so it will be zero at $t = 446.5$. But instead of getting more positive at these instants, the voltage will be *negative-going.*

If you freeze time at so-many-and-a-*quarter* seconds, say $t = 446.25$, the voltage will be +1 V. The wave will be exactly at its positive peak. If you stop time at so-many-and-*three-quarter* seconds, say $t = 446.75$, the voltage will be exactly at its negative peak, −1 V. At intermediate times, say, so-many-and-*three-tenths* seconds, the voltage will have intermediate values.

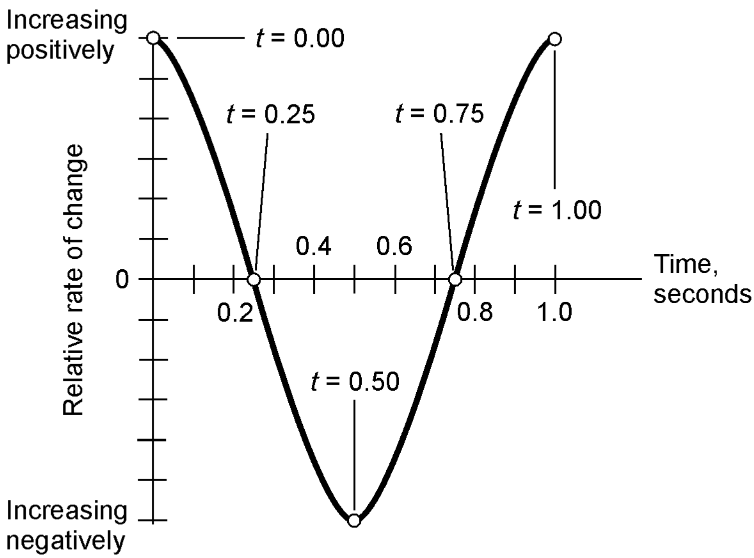
Instantaneous Rate of Change

Figure 12-1 shows that there are times the voltage is increasing, and times it is decreasing. *Increasing*, in this context, means "getting more positive," and *decreasing* means "getting more negative." The most rapid increase in voltage occurs when $t = 0.0$ and $t = 1.0$. The most rapid decrease takes place when $t = 0.5$.

When $t = 0.25$, and also when $t = 0.75$, the instantaneous voltage neither increases nor decreases. But this condition exists only for a vanishingly small moment, a single point in time.

Suppose n is some whole number. Then the situation at $t = n.25$ is the same as it is for $t = 0.25$; also, for $t = n.75$, things are the same as they are when $t = 0.75$. The single cycle shown in Fig. 12-1 represents every possible condition of the ac sine wave having a frequency of 1 Hz and a peak value of ±1 V. The whole wave recurs, over and over, for as long as the ac continues to flow in the circuit.

Now imagine that you want to observe the *instantaneous rate of change* in the voltage of the wave in Fig. 12-1, as a function of time. A graph of this turns out to be a sine wave, too—but it is displaced to the left of the original wave by ¼ of a cycle. If you plot the instantaneous rate of change of a sine wave against time (Fig. 12-2), you get the *derivative* of the waveform. The derivative of a sine wave is a *cosine wave.* This wave has the same shape as the sine wave, but the *phase* is different by ¼ of a cycle.



12-2 A sine wave representing the rate of change in the instantaneous voltage of the wave shown in Fig. 12-1.

Circles and Vectors

An ac sine wave represents the most efficient possible way that an electrical quantity can alternate. It has only one frequency component. All the wave energy is concentrated into this smoothly see-sawing variation. It is like a pure musical note.

Circular Motion

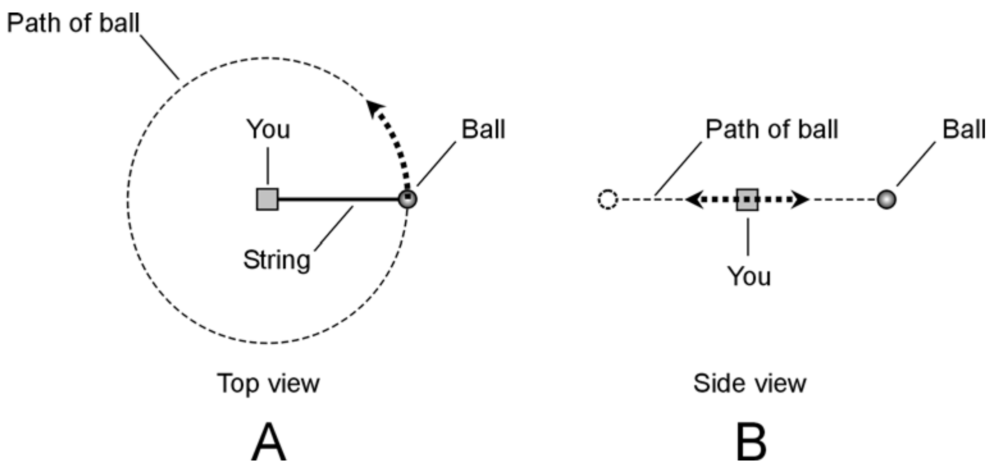
Suppose that you swing a ball around and around at the end of a string, at a rate of one revolution per second (1 rps). The ball describes a circle in space (Fig. 12-3A). If a friend stands some distance away, with his or her eyes in the plane of the ball's path, your friend sees the ball oscillating back and forth (Fig. 12-3B) with a frequency of 1 Hz. That is one complete cycle per second, because you swing the ball around at 1 rps.

If you graph the position of the ball, as seen by your friend, with respect to time, the result is a sine wave (Fig. 12-4). This wave has the same fundamental shape as all sine waves. Some sine waves are taller than others, and some are stretched out horizontally more than others. But the general waveform is the same in every case. By multiplying or dividing the amplitude and the wavelength of any sine wave, it can be made to fit exactly along the curve of any other sine wave. The standard sine wave is the function $y = \sin x$ in the coordinate plane.

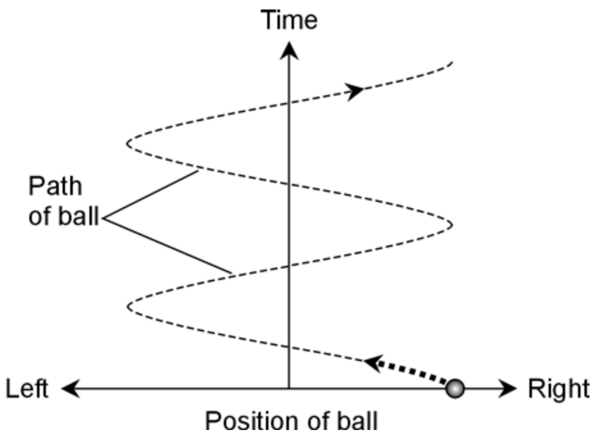
You might whirl the ball around faster or slower than 1 rps. The string might be made longer or shorter. This would alter the height and/or the frequency of the sine wave graphed in Fig. 12-4. But the sine wave can always be reduced to the equivalent of constant, smooth motion in a circular orbit. This is known as the *circular motion model* of a sine wave.

Rotating Vectors

Back in Chapter 9, *degrees of phase* were discussed. If you wondered then why phase is spoken of in terms of angular measure, the reason should be clearer now. A circle has 360°. A sine wave can be represented as circular motion. Points along a sine wave thus correspond to angles, or positions, around a circle.



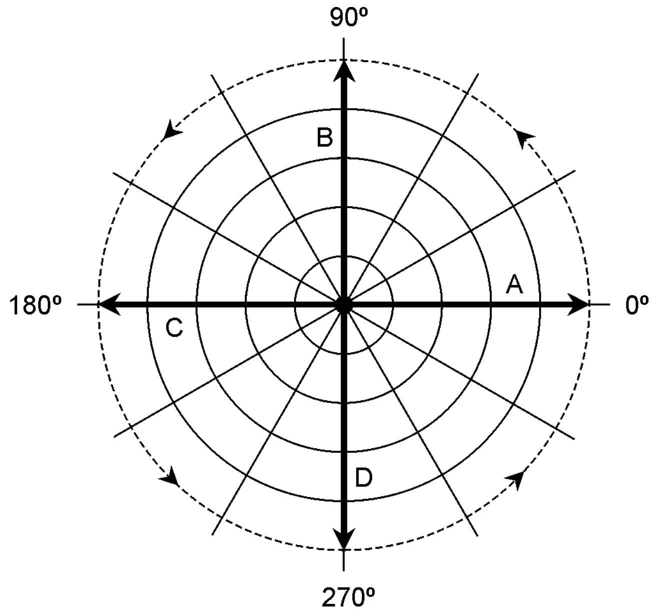
12-3 Swinging ball and string as seen from above (A) and from the side (B).



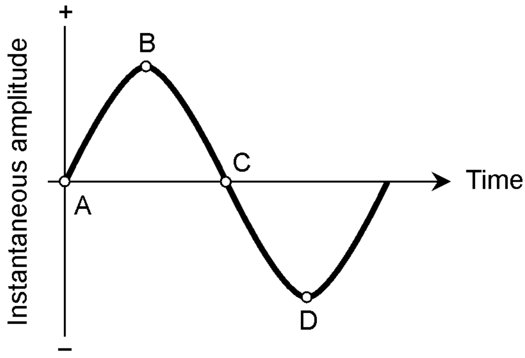
12-4 Position of ball (horizontal axis) as seen from the side, graphed as a function of time (vertical axis).

Figure 12-5 shows the way a rotating *vector* can be used to represent a sine wave. A vector is a quantity with two independent properties, called *magnitude* (or amplitude) and *direction.* At A, the vector points east, and this is assigned the value of 0°, where the wave amplitude is zero and is increasing positively. At B, the vector points north; this is the 90° instant, where the wave has attained its maximum positive amplitude. At C, the vector points west. This is 180°, the instant where the wave has gone back to zero amplitude and is getting more negative. At D, the wave points south. This is 270°, and it represents the maximum negative amplitude. When a full circle (360°) has been completed, the vector once again points east.

The four points in Fig. 12-5 are shown on a sine wave graph in Fig. 12-6. Think of the vector as revolving counterclockwise at a rate that corresponds to one revolution per cycle of the wave. If the wave has a frequency of 1 Hz, the vector goes around at a rate of 1 rps. If the wave has a frequency of



12-5 Rotating-vector representation of a sine wave. At A, at the start of the cycle; at B, one-fourth of the way through the cycle; at C, halfway through the cycle; at D, three-fourths of the way through the cycle.



12-6 The four points for the vector model of Fig. 12-5, shown in the standard amplitude-versus-time graphical manner.

100 Hz, the speed of the vector is 100 rps, or a revolution every 0.01 s. If the wave is 1 MHz, then the speed of the vector is 1 million rps (10^6 rps), and it goes once around every 0.000001 s (10^{-6} s).

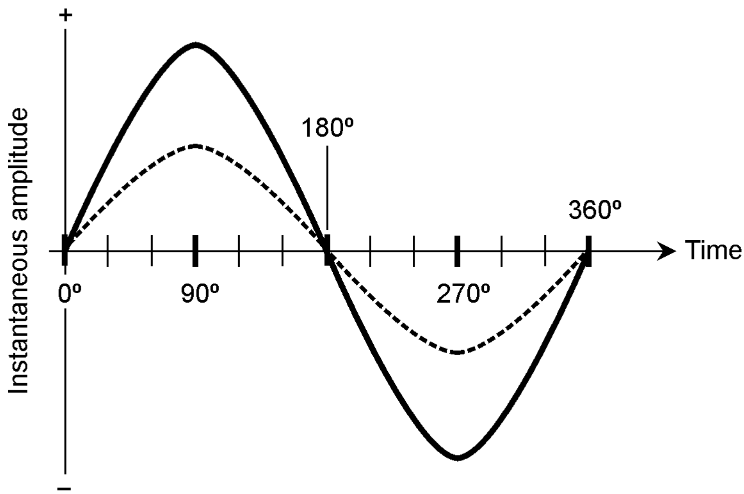
The peak amplitude of a pure ac sine wave corresponds to the length of its vector. In Fig. 12-5, time is shown by the angle counterclockwise from due east. Amplitude is independent of time. The vector length never changes, but its direction does.

Expressions of Phase Difference

The *phase difference*, also called the *phase angle*, between two waves can have meaning only when those two waves have identical frequencies. If the frequencies differ, even by just a little bit, the relative phase constantly changes, and it's impossible to specify a value for it. In the following discussions of phase angle, let's assume that the two waves always have identical frequencies.

Phase Coincidence

Phase coincidence means that two waves begin at exactly the same moment. They are "lined up." This is shown in Fig. 12-7 for two waves having different amplitudes. The phase difference in this



12-7 Two sine waves in phase coincidence.

case is 0°. You could say it's some whole-number multiple of 360°, too—but engineers and technicians rarely speak of any phase angle of less than 0° or more than 360°.

If two sine waves are in phase coincidence, and if neither wave has dc superimposed, then the resultant is a sine wave with positive or negative peak amplitudes equal to the sum of the positive and negative peak amplitudes of the composite waves. The phase of the resultant is the same as that of the composite waves.

Phase Opposition

When two sine waves begin exactly 1/2 cycle, or 180°, apart, they are said to be in *phase opposition.* This is illustrated by the drawing of Fig. 12-8. In this situation, engineers sometimes say that the waves are *out of phase*, although this expression is a little nebulous because it could be taken to mean some phase difference other than 180°.

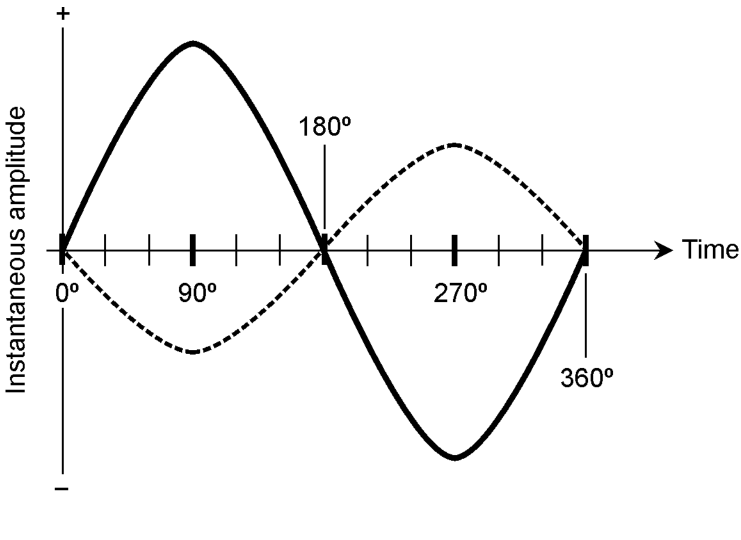
If two sine waves have the same amplitudes and are in phase opposition, they cancel each other out. This is because the instantaneous amplitudes of the two waves are equal and opposite at every moment in time.

If two sine waves are in phase opposition, and if neither wave has dc superimposed, then the resultant is a sine wave with positive or negative peak amplitudes equal to the difference between the positive and negative peak amplitudes of the composite waves. The phase of the resultant is the same as the phase of the stronger of the two composite waves.

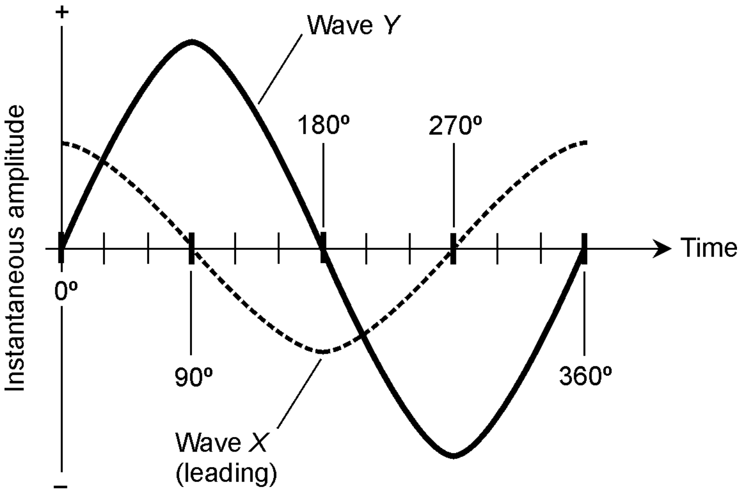
Any sine wave without superimposed dc has the unique property that, if its phase is shifted by 180°, the resultant wave is the same as turning the original wave upside down. Not all waveforms have this property. Perfect square waves do, but some rectangular and sawtooth waves don't, and irregular waveforms almost never do.

Intermediate Phase Differences

Two sine waves can differ in phase by any amount from 0° (phase coincidence), through 90° (*phase quadrature*, meaning a difference a quarter of a cycle), 180° (phase opposition), 270° (phase quadrature again), to 360° (phase coincidence again).



12-8 Two sine waves in phase opposition.



12-9 Wave X leads wave Y by 90° of phase (¼ of a cycle).

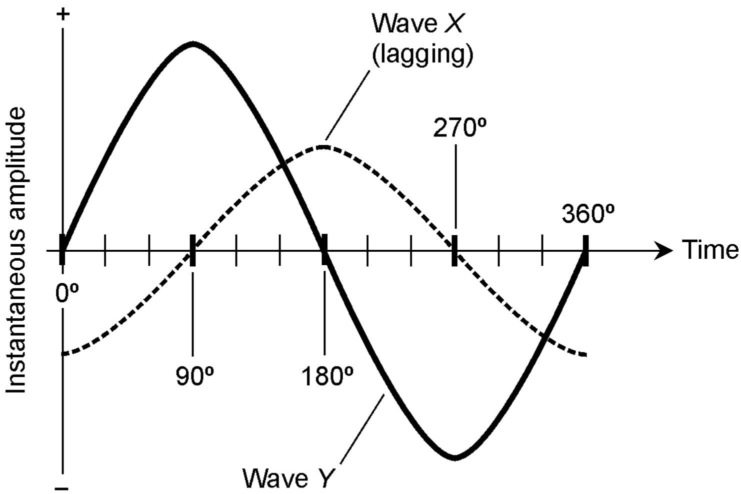
Leading Phase

Imagine two sine waves, called wave X and wave Y, with identical frequency. If wave X begins a fraction of a cycle *earlier* than wave Y, then wave X is said to be *leading* wave Y in phase. For this to be true, X must begin its cycle less than 180° before Y. Figure 12-9 shows wave X leading wave Y by 90°.

Note that if wave X (the dashed line in Fig. 12-9) is leading wave Y (the solid line), then wave X is displaced to the *left* of wave Y. In a time-domain graph or display, displacement to the left represents earlier moments in time, and displacement to the right represents later moments in time.

Lagging Phase

Suppose that some sine wave X begins its cycle more than 180°, but less than 360°, ahead of wave Y. In this situation, it is easier to imagine that wave X starts its cycle *later* than wave Y, by some value between 0° and 180°. Then wave X is not leading, but *lagging*, wave Y. Figure 12-10 shows wave X lagging wave Y by 90°.



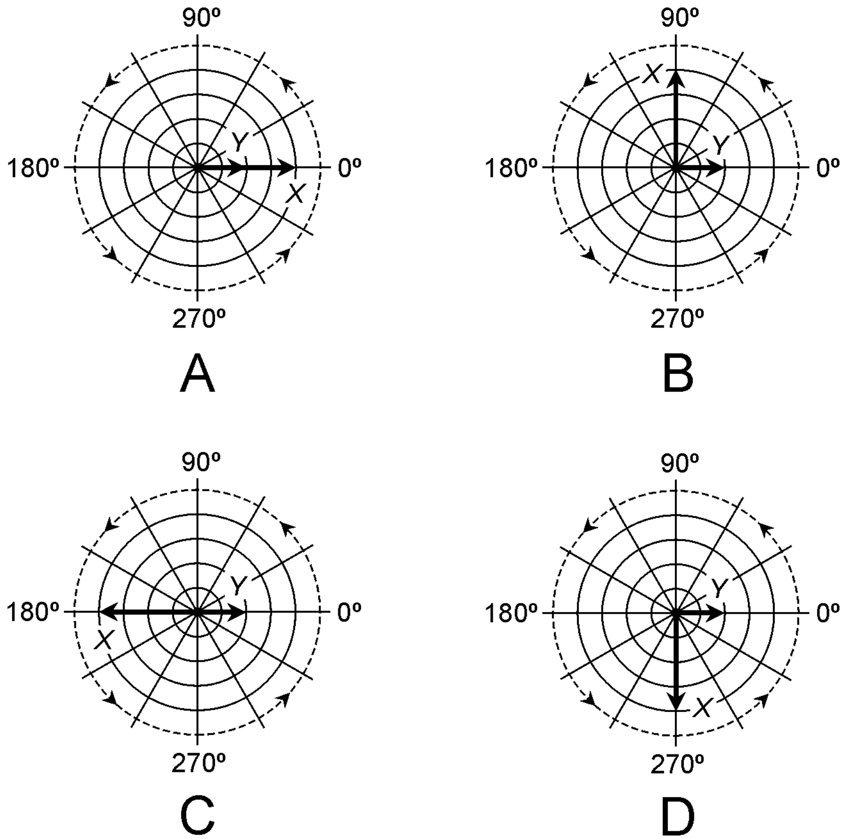
12-10 Wave X lags wave Y by 90° of phase (¼ of a cycle).

Vector Diagrams of Phase Difference

The vector renditions of sine waves, such as are shown in Fig. 12-5, are well suited to showing phase relationships.

If a sine wave X leads a sine wave Y by some number of degrees, then the two waves can be drawn as vectors, with vector **X** being that number of angular degrees *counterclockwise* from vector **Y.** If a sine wave X lags a sine wave Y by some number of degrees, then **X** appears to point in a direction that is *clockwise* from **Y** by that number of angular degrees. If two waves are in phase coincidence, then their vectors point in exactly the same direction. If two waves are in phase opposition, then their vectors point in exactly opposite directions.

The drawings of Fig. 12-11 show four phase relationships between two sine waves X and Y. At A, X is in phase with Y. At B, X leads Y by 90°. At C, X and Y are 180° apart in phase. At D, X lags Y by 90°. In all of these examples, think of the vectors rotating *counterclockwise* as time passes, but always maintaining the same angle with respect to each other, and always staying at the same lengths. If the frequency in hertz is f, then the pair of vectors rotates together, counterclockwise, at an angular speed of f, expressed in complete 360° revolutions per second.



12-11 Vector representations of phase difference. At A, waves X and Y are in phase. At B, X leads Y by 90°. At C, X and Y are 180° out of phase. At D, X lags Y by 90°. Time is represented by counterclockwise motion of both vectors at a constant angular speed.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. Which of the following is *not* a general characteristic of an ac wave?
   (a) The wave shape is identical for each cycle.
   (b) The polarity reverses periodically.
   (c) The electrons always flow in the same direction.
   (d) There is a definite frequency.

2. All sine waves
   (a) have similar general appearance.
   (b) have instantaneous rise and fall times.
   (c) are in the same phase as cosine waves.
   (d) rise instantly, but decay slowly.

3. The derivative of a sine wave
   (a) is shifted in phase by ½ cycle from the sine wave.
   (b) is the rate of change in the instantaneous value.
   (c) has instantaneous rise and decay times.
   (d) rises instantly, but decays slowly.

4. A phase difference of 180° in the circular motion model of a sine wave represents
   (a) ¼ revolution.
   (b) ½ revolution.
   (c) a full revolution.
   (d) two full revolutions.

5. You can add or subtract a certain number of degrees of phase to or from a wave, and end up with exactly the same wave again. This number is
   (a) 90, or any whole-number multiple of it.
   (b) 180, or any whole-number multiple of it.
   (c) 270, or any whole-number multiple of it.
   (d) 360, or any whole-number multiple of it.

6. You can add or subtract a certain number of degrees of phase to or from a sine wave, and end up with an inverted (upside-down) representation of the original. This number is
   (a) 90, or any odd whole-number multiple of it.
   (b) 180, or any odd whole-number multiple of it.
   (c) 270, or any odd whole-number multiple of it.
   (d) 360, or any odd whole-number multiple of it.

7. Suppose a wave has a frequency of 300 kHz. How long does one complete cycle take?
   (a) 1.300 s
   (b) 0.00333 s
   (c) 1/3000 s
   (d) 3.33×10^{-6} s

8. If a wave has a frequency of 440 Hz, how long does it take for 10° of a cycle to occur?
   (a) 0.00273 s
   (b) 0.000273 s
   (c) 0.0000631 s
   (d) 0.00000631 s

9. Suppose two waves are in phase coincidence. One has peak values of ±3 V and the other has peak values of ±5 V. The resultant has voltages of
   (a) ±8 V pk, in phase with the composites.
   (b) ±2 V pk, in phase with the composites.
   (c) ±8 V pk, in phase opposition with respect to the composites.
   (d) ±2 V pk, in phase opposition with respect to the composites.

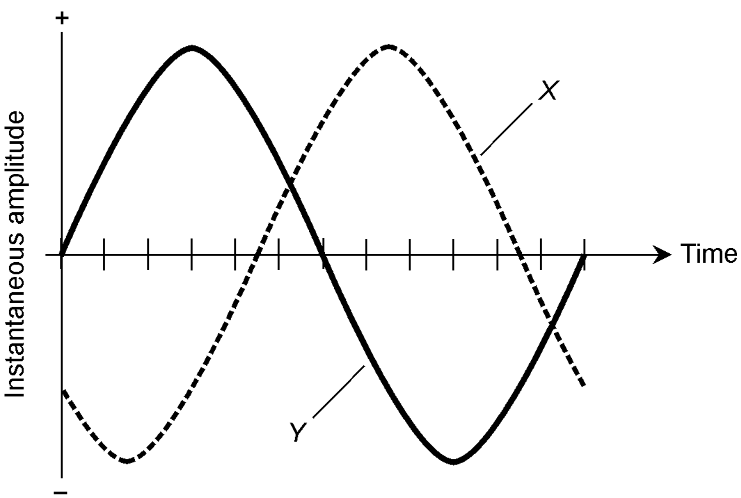
10. As shown on a graph, shifting the phase of an ac sine wave by 90° is the same thing as
   (a) moving it to the right or left by a full cycle.
   (b) moving it to the right or left by ¼ cycle.
   (c) turning it upside down.
   (d) leaving it alone.

11. Two pure sine waves that differ in phase by 180° can be considered to
   (a) be offset by two full cycles.
   (b) be in phase opposition.
   (c) be separated by less than ¼ cycle.
   (d) have a frequency of ½ cycle.

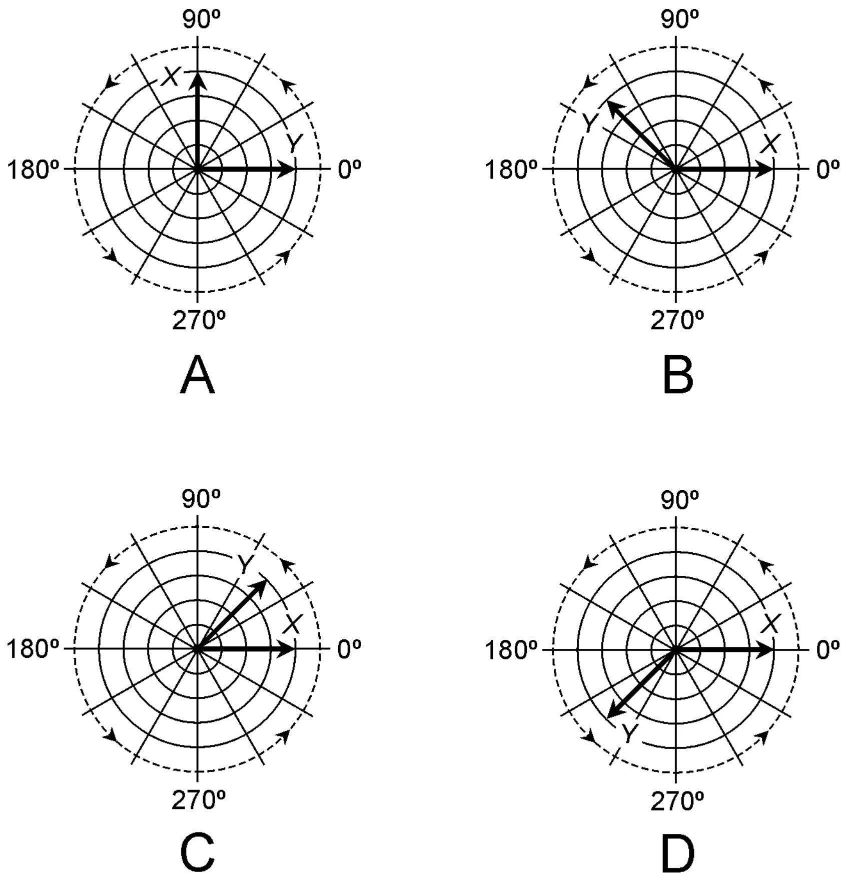
12. Suppose two sine waves are in phase opposition. Wave *X* has a peak amplitude of ±4 V and wave *Y* has a peak amplitude of ±8 V. The resultant has voltages of
   (a) ±4 V pk, in phase with the composites.
   (b) ±4 V pk, out of phase with the composites.
   (c) ±4 V pk, in phase with wave *X*.
   (d) ±4 V pk, in phase with wave *Y*.

13. If wave *X* leads wave *Y* by 45°, then
   (a) wave *Y* is ¼ cycle ahead of wave *X*.
   (b) wave *Y* is ¼ cycle behind wave *X*.
   (c) wave *Y* is ⅛ cycle behind wave *X*.
   (d) wave *Y* is 1.16 cycle ahead of wave *X*.

14. If wave *X* lags wave *Y* by ⅓ cycle, then
    (a) wave *Y* is 120° ahead of wave *X*.
    (b) wave *Y* is 90° ahead of wave *X*.
    (c) wave *Y* is 60° ahead of wave *X*.
    (d) wave *Y* is 30° ahead of wave *X*.
15. Refer to Fig. 12-12. In this example,
    (a) *X* lags *Y* by 45°.
    (b) *X* leads *Y* by 45°.
    (c) *X* lags *Y* by 135°.
    (d) *X* leads *Y* by 135°.
16. Which of the drawings in Fig. 12-13 represents the situation of Fig. 12-12?
    (a) Drawing A
    (b) Drawing B
    (c) Drawing C
    (d) Drawing D
17. In vector diagrams such as those of Fig. 12-13, the length of the vector represents
    (a) the average amplitude of a sine wave.
    (b) the frequency of a sine wave.
    (c) the phase of a sine wave.
    (d) the peak amplitude of a sine wave.
18. In vector diagrams such as those of Fig. 12-13, the angle between two vectors represents
    (a) the average of the peak amplitudes of two sine waves.
    (b) the frequency difference between two sine waves.
    (c) the phase difference between two sine waves.
    (d) the difference between the peak amplitudes of two sine waves.



12-12 Illustration for Quiz Question 15.



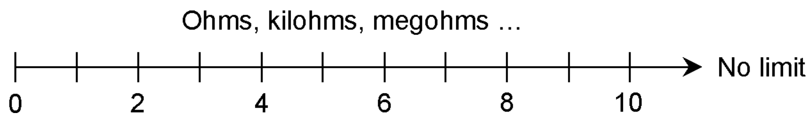
12-13 Illustration for Quiz Questions 16 through 20.

19. In vector diagrams such as those of Fig. 12-13, the distance from the center of the graph represents
    (a) average amplitude.
    (b) frequency.
    (c) phase.
    (d) peak amplitude.
20. In diagrams like those of Fig. 12-13, the progression of time is sometimes depicted as
    (a) movement of a vector to the right.
    (b) movement of a vector to the left.
    (c) counterclockwise rotation of a vector.
    (d) clockwise rotation of a vector.

13 CHAPTER

Inductive Reactance

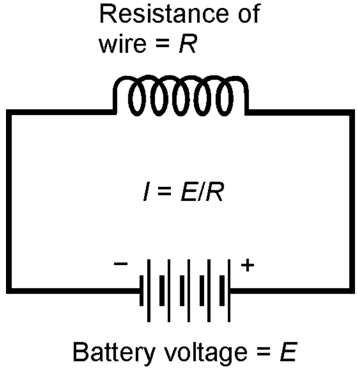
IN DC CIRCUITS, RESISTANCE CAN BE EXPRESSED AS A NUMBER RANGING FROM ZERO (REPRESENTING a perfect conductor) to extremely large values. Physicists call resistance a *scalar* quantity, because it can be expressed on a one-dimensional *scale*, as shown in Fig. 13-1.



13-1 Resistance can be represented as numerical values (corresponding to ohms) along a half line or ray.

Coils and Direct Current

Suppose you have some wire that conducts electricity very well. If you wind a length of the wire into a coil and connect it to a source of dc (Fig. 13-2), the wire draws a large current. It doesn't matter



13-2 An inductor connected across a source of dc.

whether the wire is a single-turn loop, or whether it's lying haphazardly on the floor, or whether it's wrapped around a stick. The current amperes is equal to the applied voltage in volts divided by the wire resistance in ohms. It's that simple.

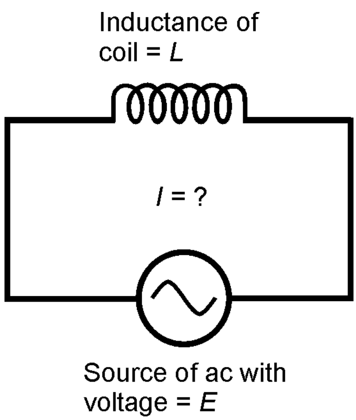
You can make an electromagnet, as you've already seen, by passing dc through a coil wound around an iron rod. Electromagnets are known for the high current they draw from batteries or power supplies. The coil of an electromagnet heats up as energy is dissipated in the resistance of the wire. If the voltage of the battery or power supply increases, the wire in the coil gets hotter. Ultimately, if the supply can deliver enough current, the wire will melt.

Coils and Alternating Current

Suppose you change the voltage source, connected across the coil, from dc to ac (Fig. 13-3). Imagine that you can vary the frequency of the ac, from a few hertz to hundreds of hertz, then kilohertz, then megahertz.

At first, the current will be high, just as it is with dc. But the coil has a certain amount of inductance, and it takes some time for current to establish itself in the coil. Depending on how many turns there are and on whether the core is air or a ferromagnetic material, you'll reach a point, as the ac frequency increases, when the coil starts to get sluggish. That is, the current won't have time to get established in the coil before the polarity of the ac voltage reverses. At high ac frequencies, the current through the coil will have difficulty following the voltage placed across the coil. This sluggishness in a coil for ac is, in effect, similar to dc resistance. As the frequency is raised, the effect gets more pronounced. Eventually, if you keep increasing the frequency of the ac source, the coil will not even come near establishing a current with each cycle. Then the coil will act like a high resistance.

The opposition that the coil offers to ac is called *inductive reactance.* It, like resistance, is measured in ohms. It can vary, just as resistance does, from near zero (a short piece of wire) to a few ohms (a small coil) to kilohms or megohms (bigger and bigger coils). Like resistance, inductive reactance affects the current in an ac circuit. But, unlike simple resistance, reactance changes with frequency. This effect is not merely a decrease in the current, although in practice this does happen. Inductive reactance produces a change in the way the current flows with respect to the voltage.



13-3 An inductor connected across a source of ac.

Reactance and Frequency

Inductive reactance is one of two form of reactance. (The other form, called *capacitive reactance*, will be discussed in the next chapter.) Reactance in general is symbolized by the italic uppercase letter X. Inductive reactance is symbolized X_L.

If the frequency of an ac source is given, in hertz, as f, and the inductance of a coil in henrys is given as L, then the inductive reactance in ohms, X_L, is calculated as follows:

$$X_L = 2\pi fL$$

In this formula, the symbol π stands for the mathematical constant *pi*, which is the number of diameters around the circumference of a circle. It is equal to approximately 3.14. We can consider the value of 2π to be equal to 6.28 in most practical situations. Therefore, the preceding formula can be written a little more simply as:

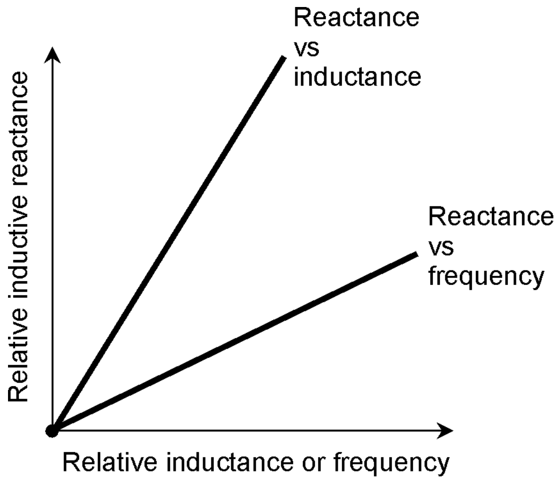
$$X_L = 6.28fL$$

This same formula applies if the frequency, f, is in kilohertz and the inductance, L, is in millihenrys. And it also applies if f is in megahertz and L is in microhenrys. Just remember that if frequency is in thousands, inductance must be in thousandths, and if frequency is in millions, inductance must be in millionths.

Inductive reactance increases *linearly* with increasing ac frequency. This means that the function of X_L versus f is a straight line when graphed. Inductive reactance also increases linearly with inductance. Therefore, the function of X_L versus L also appears as a straight line on a graph. The value of X_L is *directly proportional* to f, and is also directly proportional to L. These relationships are graphed, in relative form, in Fig. 13-4.

Problem 13-1

Suppose a coil has an inductance of 0.500 H, and the frequency of the ac passing through it is 60.0 Hz. What is the inductive reactance?



13-4 Inductive reactance is directly proportional to inductance, and is also directly proportional to frequency.

Using the preceding formula, calculate $X_L = 6.28 \times 60.0 \times 0.500 = 188\ \Omega$. This is rounded to three significant figures.

Problem 13-2

What will be the inductive reactance of the preceding coil if the supply is a battery that supplies pure dc?

Because dc has a frequency of zero, $X_L = 6.28 \times 0 \times 0.500 = 0\ \Omega$. That is, there will be no inductive reactance. Inductance doesn't have any practical effect with pure dc.

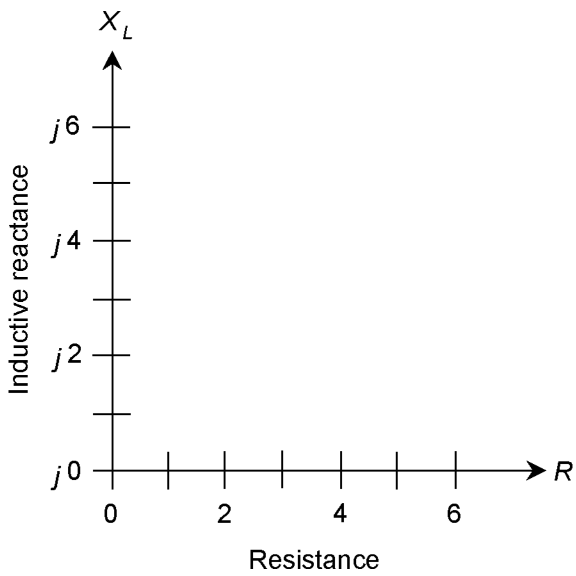
Problem 13-3

If a coil has an inductive reactance of 100 Ω at a frequency of 5.00 MHz, what is its inductance?

In this case, you need to plug numbers into the formula and solve for the unknown L. Start out with the equation $100 = 6.28 \times 5.00 \times L = 31.4 \times L$. Because the frequency is given in megahertz, the inductance will come out in microhenrys. You can divide both sides of the equation by 31.4, getting $L = 100/31.4 = 3.18\ \mu H$.

Points in the *RL* Plane

Inductive reactance can be plotted along a half line, just as can resistance. In a circuit containing both resistance and inductance, the characteristics become two-dimensional. You can orient the resistance and reactance half lines perpendicular to each other to make a quarter-plane coordinate system, as shown in Fig. 13-5. Resistance is plotted horizontally, and inductive reactance is plotted vertically upward.



13-5 The quarter plane for inductive reactance (X_L) and resistance (R). This is also known as the *RL* quarter-plane, or simply as the *RL* plane.

In this scheme, resistance-inductance (*RL*) combinations form *complex impedances.* (The term *impedance* comes from the root *impede*, and fully describes how electrical components impede, or inhibit, the flow of ac. You'll learn all about this in Chap. 15.) Each point on the *RL plane* corresponds to one unique complex impedance value. Conversely, each complex impedance value corresponds to one unique point on the *RL* plane.

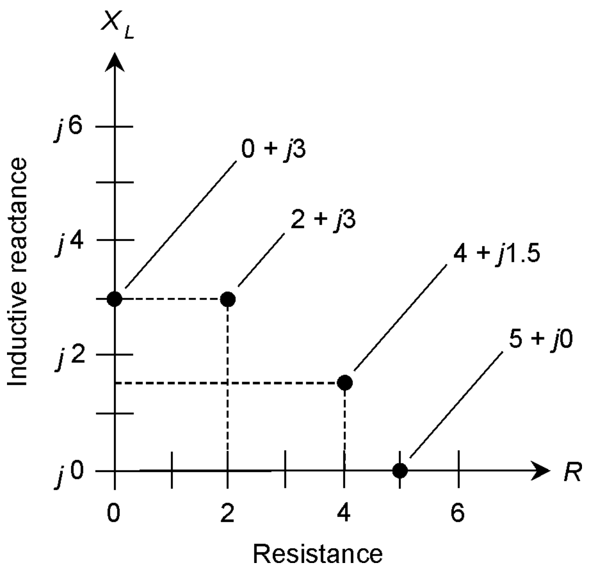
You might ask, "What's the little *j* doing in Fig. 13-5?" For reasons that will be made clear in Chap. 15, impedances on the *RL* plane are written in the form $R + jX_L$, where R is the resistance in ohms, and X_L is the inductive reactance in ohms. The little *j* is called a *j operator* and is a mathematical way of expressing the fact that reactance is denoted at right angles to resistance in complex-impedance graphs.

If you have a pure resistance, say $R = 5\ \Omega$, then the *complex* impedance is $5 + j0$, and is at the point (5,0) on the *RL* plane. If you have a pure inductive reactance, such as $X_L = 3\ \Omega$, then the complex impedance is $0 + j3$, and is at the point $(0,j3)$ on the *RL* plane. These points, and a couple of others, are shown in Fig. 13-6.

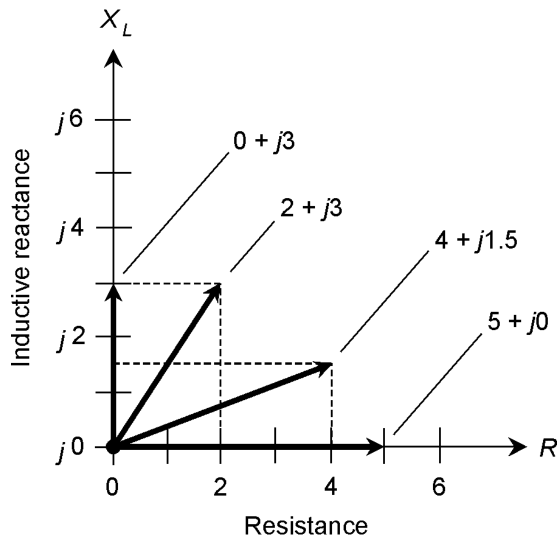
In real life, all coils have some resistance, because no wire is a perfect conductor. All resistors have at least a tiny bit of inductive reactance, because they take up some physical space and they have wire leads. So there is really no such thing as a mathematically perfect pure resistance such as $5 + j0$, or a mathematically perfect pure reactance like $0 + j3$. But sometimes you can get extremely close to theoretical ideals in real life.

Often, resistance and inductive reactance are both deliberately placed in a circuit. Then you get impedances values such as $2 + j3$ or $4 + j1.5$. These are shown in Fig. 13-6 as points on the *RL* plane.

Remember that values for X_L are *reactances*, not actual inductances. Because of this, they vary with the frequency in an *RL* circuit. Changing the frequency has the effect of making complex impedance points move around in the *RL* plane. They move vertically, going upward as the ac frequency increases, and downward as the ac frequency decreases. If the ac frequency goes down to zero, the inductive reactance vanishes. Then $X_L = 0$, we have pure dc, and the point is right on the resistance axis.



13-6 Four points in the *RL* plane.



13-7 Four vectors in the *RL* plane, corresponding to the points shown in Fig. 13-6.

Vectors in the *RL* Plane

Engineers sometimes represent points in the *RL* plane as vectors. Recall that a vector is a mathematical quantity that has a defined magnitude (length) and defined direction (orientation). Expressing a point in the *RL* plane as a vector thus gives that point a unique magnitude and a unique direction.

In Fig. 13-6, four different points are shown. Each point is represented by a certain distance to the right of the *origin* $(0,j0)$, and a certain distance upward from the origin. The first of these is the resistance, R, and the second is the inductive reactance, X_L. Thus, the *RL* combination is a two-dimensional quantity. There is no way to uniquely define *RL* combinations as single numbers, or *scalars*, because there are two different quantities that can vary independently.

Another way to depict these points is to draw lines from the origin out to them. Then you can think of the points as *rays*, each having a certain length, or magnitude, and a certain direction, or angle counterclockwise from the resistance axis. These rays, going out to the points, are *complex impedance vectors* (Fig. 13-7).

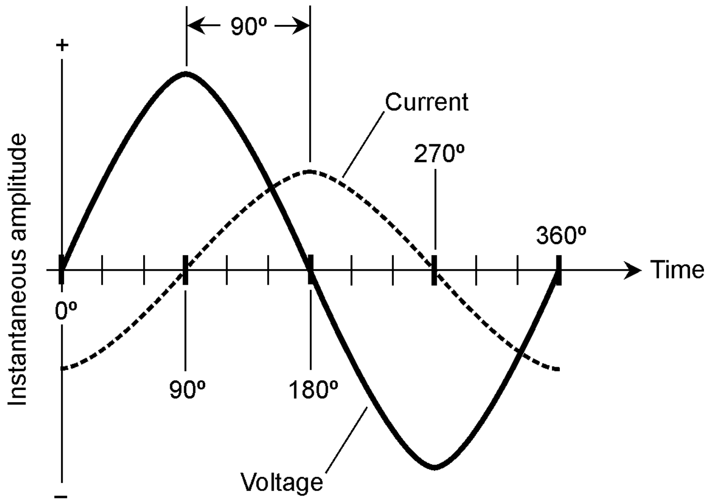
Current Lags Voltage

When an ac voltage is placed across an inductor and starts to increase (either positively or negatively) from zero, it takes a fraction of a cycle for the current to follow. Once the voltage starts decreasing from its maximum peak (either positive or negative) in the cycle, it again takes a fraction of a cycle for the current to follow. The instantaneous current can't quite keep up with the instantaneous voltage, as it does in a pure resistance. Thus, in a circuit containing inductive reactance, the current is said to *lag* the voltage in phase.

Pure Inductance

Suppose that you place an ac voltage across a coil, with a frequency high enough so that the inductive reactance, X_L, is much larger than the resistance, R. In this situation, the current is $\frac{1}{4}$ of a cycle behind the voltage. That is, the current lags the voltage by 90°, as shown in Fig. 13-8.

At very low frequencies, large inductances are normally needed in order for the current lag to be



13-8 In a pure inductive reactance, the current lags the voltage by 90°.

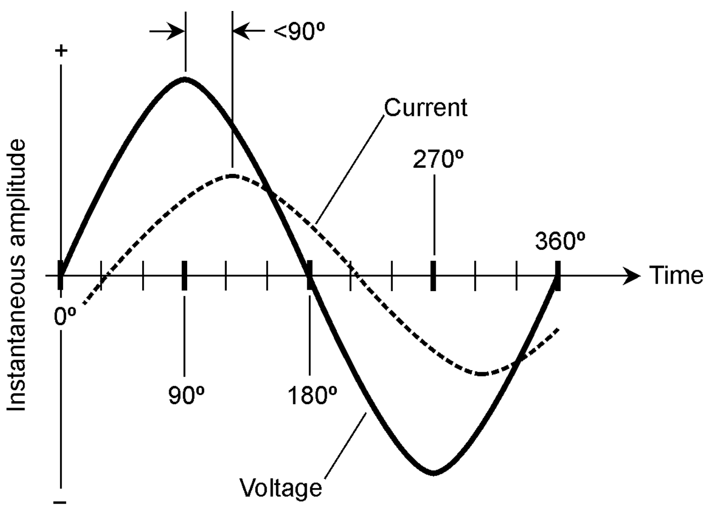
a full 90°. This is because any coil has some resistance; no wire is a perfect conductor. If some wire were found that had a mathematically zero resistance, and if a coil of any size were wound from this wire, then the current would lag the voltage by 90° in this inductor, no matter what the ac frequency.

When the value of X_L is very large compared with the value of R in a circuit—that is, when there is an essentially pure inductive reactance—the vector in the RL plane points straight up along the X_L axis. Its angle is 90° from the R axis, which is considered the *zero line* in the RL plane.

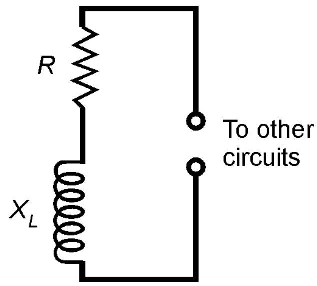
Inductance with Resistance

When the resistance in a resistance-inductance (RL) circuit is significant compared with the inductive reactance, the current lags the voltage by something less than 90° (Fig. 13-9). If R is small compared with X_L, the current lag is almost 90°, but as R gets larger relative to X_L, the lag decreases.

The value of R in an RL circuit can increase relative to X_L because resistance is deliberately placed in series with the inductance. It can also happen because the ac frequency gets so low that X_L



13-9 In a circuit with inductive reactance and resistance, the current lags the voltage by less than 90°.



13-10 Schematic representation of a circuit containing resistance and inductive reactance.

decreases until it is comparable to the loss resistance R in the coil winding. In either case, the situation can be schematically represented by an inductance in series with a resistance (Fig. 13-10).

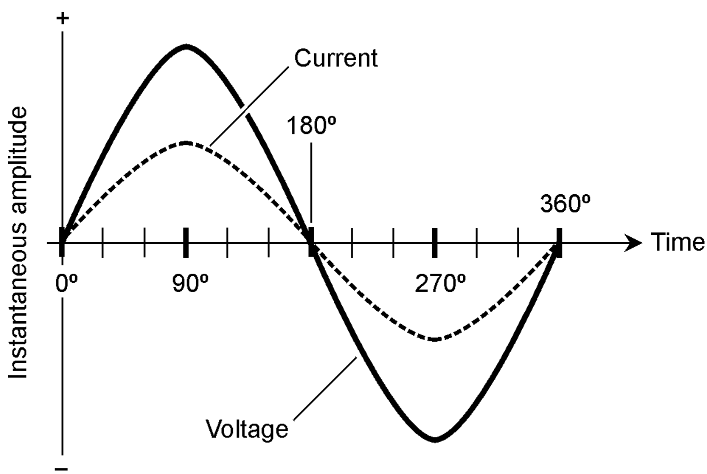
If you know the values of X_L and R, you can find the *angle of lag*, also called the *RL phase angle*, by plotting the point $R + jX_L$ on the *RL* plane, drawing the vector from the origin out to that point, and then measuring the angle of the vector, counterclockwise from the resistance axis. You can use a protractor to measure this angle, or you can compute its value using trigonometry.

Actually, you don't have to know the actual values of X_L and R in order to find the angle of lag. All you need to know is their ratio. For example, if $X_L = 5\ \Omega$ and $R = 3\ \Omega$, you get the same *RL* phase angle that you get if $X_L = 50\ \Omega$ and $R = 30\ \Omega$, or if $X_L = 20\ \Omega$ and $R = 12\ \Omega$. The angle of lag is the same for any values of X_L and R in the ratio 5:3.

Pure Resistance

As the resistance in an *RL* circuit becomes large with respect to the inductive reactance, the angle of lag gets small. The same thing happens if the inductive reactance gets small compared with the resistance. When R is many times greater than X_L, the vector in the *RL* plane lies almost on the R axis, going east (to the right). The *RL* phase angle in this case is close to 0°. The current is nearly in phase with the voltage.

In a pure resistance, with no inductance at all, the current is precisely in phase with the voltage (Fig. 13-11). A pure resistance doesn't store and release energy as an inductive circuit does, so there is no sluggishness in it.



13-11 In a circuit with pure resistance (no reactance), the current is in phase with the voltage.

How Much Lag?

If you know the ratio of the inductive reactance to the resistance (X_L/R) in an *RL* circuit, then you can find the phase angle. Of course, you can also find the phase angle if you know the actual values of X_L and R.

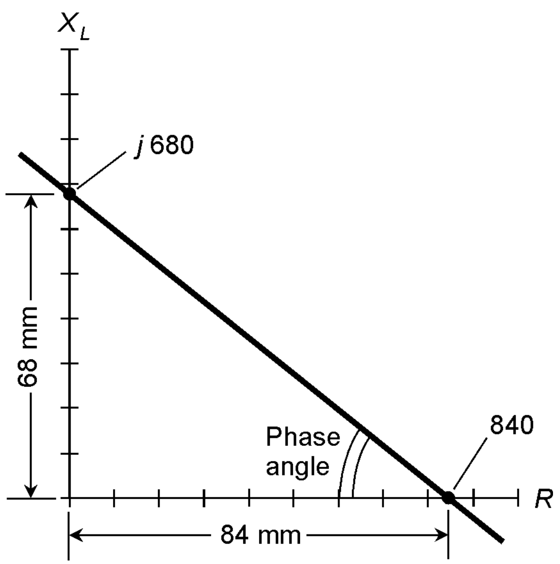
Pictorial Method

It isn't necessary to construct an entire *RL* plane to find phase angles. You can use a ruler that has centimeter (cm) and millimeter (mm) markings, and a protractor. First, draw a line a little more than 10 cm long, going from left to right on a sheet of paper. Use the ruler and a sharp pencil. Then, with the protractor, construct a line off the left end of this first line, going vertically upward. Make this line at least 10 cm long. The horizontal line, or the one going to the right, is the R axis of a coordinate system. The vertical line, or the one going upward, is the X_L axis.

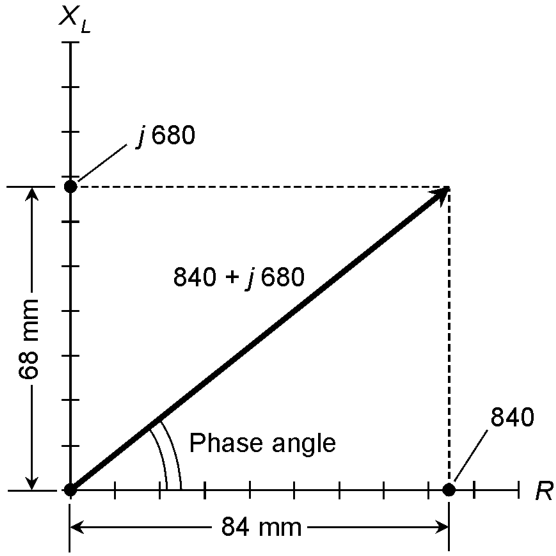
If you know the values of X_L and R, divide them down or multiply them up so they're both between 0 and 100. For example, if $X_L = 680\ \Omega$ and $R = 840\ \Omega$, you can divide them both by 10 to get $X_L = 68$ and $R = 84$. Plot these points lightly by making hash marks on the vertical and horizontal lines you've drawn. The R mark in this example will be 84 mm to the right of the origin, and the X_L mark will be 68 mm up from the origin.

Next, draw a line connecting the two hash marks, as shown in Fig. 13-12. This line will run at a slant, and will form a triangle along with the two axes. Your hash marks, and the origin of the coordinate system, form the three *vertices* of a *right triangle.* The triangle is called *right* because one of its angles is a right angle (90°). Measure the angle between the slanted line and the R axis. Extend one or both of the lines if necessary in order to get a good reading on the protractor. This angle will be between 0 and 90°, and represents the phase angle in the *RL* circuit.

The complex impedance vector, $R + jX_L$, is found by constructing a rectangle using the origin and your two hash marks as three of the four vertices, and drawing new horizontal and vertical lines to complete the figure. The vector is the diagonal of this rectangle, as shown in Fig. 13-13. The



13-12 Pictorial method of finding phase angle in a circuit containing resistance and inductive reactance.



13-13 Another pictorial method of finding phase angle in a circuit containing resistance and inductive reactance. This method shows the actual impedance vector.

phase angle is the angle between this vector and the *R* axis. It will be the same as the angle of the slanted line in Fig. 13-12.

Trigonometric Method

If you have a good scientific calculator that can find the *arctangent* of a number (also called the *inverse tangent* and symbolized either as arctan or $\tan^{-1}$), you can determine the *RL* phase angle more precisely than the pictorial method allows. Given the values of X_L and R, the *RL* phase angle is the arctangent of their ratio. Phase angle is symbolized by the lowercase Greek letter phi (pronounced "fie" or "fee" and written ϕ). Therefore:

$$\phi = \tan^{-1}\,(X_L/R) \qquad \text{or} \qquad \phi = \arctan\,(X_L/R)$$

Problem 13-4

Suppose the inductive reactance in an *RL* circuit is 680 Ω and the resistance is 840 Ω. What is the phase angle?

The ratio X_L/R is 680/840. A calculator will display this quotient as something like 0.8095 and some more digits. Find the arctangent of this number. You should get 38.99 and some more digits. This can be rounded off to 39.0°.

Problem 13-5

Suppose an *RL* circuit operates at a frequency of 1.0 MHz with a resistance of 10 Ω and an inductance of 90 μH. What is the phase angle? What does this tell us about the nature of this *RL* circuit at this frequency?

Find the inductive reactance using the formula $X_L = 6.28fL = 6.28 \times 1.0 \times 90 = 565$ Ω. Then find the ratio $X_L/R = 565/10 = 56.5$. The phase angle is equal to arctan 56.5, which, rounded to two significant figures, is 89°. The circuit contains an almost pure inductive reactance, because the phase angle is close to 90°. The resistance contributes little to the behavior of this *RL* circuit at 1.0 MHz.

Problem 13-6

What is the phase angle for the preceding circuit at a frequency of 10 kHz? With that information, what can we say about the behavior of the circuit at 10 kHz?

This requires that X_L be calculated again, for the new frequency. Let's use megahertz, so it goes in the formula with microhenrys. A frequency of 10 kHz is the same as 0.010 MHz. Calculating, we get $X_L = 6.28fL = 6.28 \times 0.010 \times 90 = 5.65\ \Omega$. The ratio X_L/R is $5.65/10 = 0.565$. Therefore, the phase angle is arctan 0.565, which, rounded to two significant figures, is 29°. This is not close to either 0° or 90°. Thus, at 10 kHz, the resistance and the inductive reactance both play significant roles in the behavior of the circuit.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. As the number of turns in a coil that carries ac increases without limit, the current in the coil will
   (a) eventually become very large.
   (b) stay the same.
   (c) decrease, approaching zero.
   (d) be stored in the core material.

2. As the number of turns in a coil increases, the reactance at a constant frequency
   (a) increases.
   (b) decreases.
   (c) stays the same.
   (d) is stored in the core material.

3. As the frequency of an ac wave gets lower, the value of X_L for a particular coil of wire
   (a) increases.
   (b) decreases.
   (c) stays the same.
   (d) depends on the voltage.

4. Suppose a coil has an inductance of 100 mH. What is the reactance at a frequency of 1000 Hz?
   (a) 0.628 Ω
   (b) 6.28 Ω
   (c) 62.8 Ω
   (d) 628 Ω

5. Suppose a coil shows an inductive reactance of 200 Ω at 500 Hz. What is its inductance?
   (a) 0.637 H
   (b) 628 H
   (c) 63.7 mH
   (d) 628 mH
6. Imagine a 400-μH inductor with a reactance of 33 Ω. What is the frequency?
   (a) 13 kHz
   (b) 0.013 kHz
   (c) 83 kHz
   (d) 83 MHz
7. Suppose an inductor has $X_L = 555\ \Omega$ at $f = 132$ kHz. What is *L?*
   (a) 670 mH
   (b) 670 μH
   (c) 460 mH
   (d) 460 μH
8. Suppose a coil has $L = 689\ \mu$H at $f = 990$ kHz. What is X_L?
   (a) 682 Ω
   (b) 4.28 Ω
   (c) 4.28 kΩ
   (d) 4.28 MΩ
9. Suppose an inductor has $L = 88$ mH with $X_L = 100\ \Omega$. What is f?
   (a) 55.3 kHz
   (b) 55.3 Hz
   (c) 181 kHz
   (d) 181 Hz
10. Each point in the *RL* plane
    (a) corresponds to a unique resistance.
    (b) corresponds to a unique inductance.
    (c) corresponds to a unique combination of resistance and inductive reactance.
    (d) corresponds to a unique combination of resistance and inductance.
11. If the resistance R and the inductive reactance X_L both are allowed to vary from zero to unlimited values, but are always in the ratio 3:1, the points in the *RL* plane for all the resulting impedances will lie along
    (a) a vector pointing straight up.
    (b) a vector pointing east.
    (c) a circle.
    (d) a ray of indefinite length, pointing outward from the origin.

12. Each specific complex impedance value defined in the form $R + jX_L$
    (a) corresponds to a specific point in the *RL* plane.
    (b) corresponds to a specific inductive reactance.
    (c) corresponds to a specific resistance.
    (d) All of the above are true.

13. A vector is defined as a mathematical quantity that has
    (a) magnitude and direction.
    (b) resistance and inductance.
    (c) resistance and reactance.
    (d) inductance and reactance.

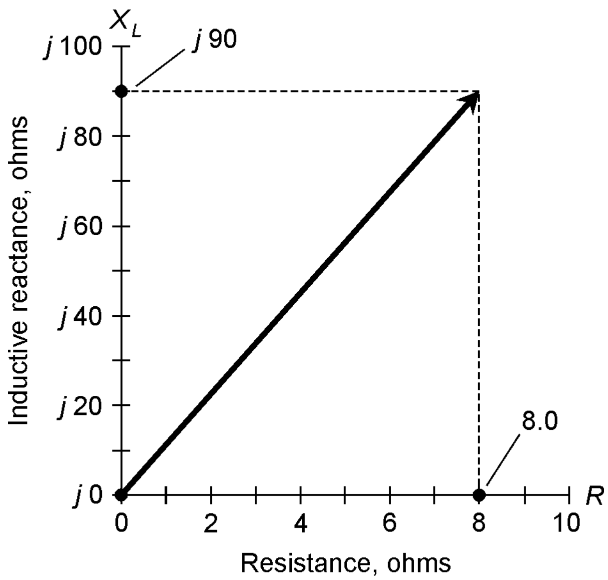
14. In an *RL* circuit, as the ratio of inductive reactance to resistance (X_L/R) decreases, the phase angle
    (a) increases.
    (b) decreases.
    (c) stays the same.
    (d) becomes alternately positive and negative.

15. In a circuit containing inductive reactance but no resistance, the phase angle is
    (a) constantly increasing.
    (b) constantly decreasing.
    (c) equal to 0°.
    (d) equal to 90°.

16. If the inductive reactance and the resistance in an *RL* circuit are equal (as expressed in ohms), then what is the phase angle?
    (a) 0°
    (b) 45°
    (c) 90°
    (d) It depends on the actual values of the resistance and the inductive reactance.

17. In Fig. 13-14, the impedance shown is which of the following?
    (a) 8.0 Ω
    (b) 90 Ω
    (c) $90 + j8.0$
    (d) $8.0 + j90$

18. Note that in the diagram of Fig. 13-14, the R and X_L scale divisions are of different sizes. The phase angle can nevertheless be determined. It is
    (a) about 50°, from the looks of it.
    (b) 48°, as measured with a protractor.



13-14 Illustration for Quiz Questions 17 and 18.

(c) 85°, as calculated using trigonometry.

(d) 6.5°, as calculated using trigonometry.

19. Consider an *RL* circuit that consists of a 100-μH inductor and a 100-Ω resistor. What is the phase angle at a frequency of 200 kHz?

(a) 45.0°

(b) 51.5°

(c) 38.5°

(d) There isn't enough data given to calculate it.

20. Suppose an *RL* circuit has an inductance of 88 mH, and the resistance is 95 Ω. At 800 Hz, what is the phase angle?

(a) 78°

(b) 12°

(c) 43°

(d) 47°

14
CHAPTER

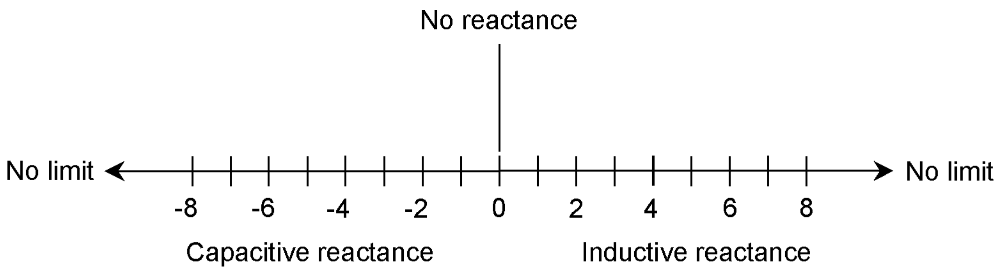
Capacitive Reactance

CAPACITIVE REACTANCE IS THE NATURAL COUNTERPART OF INDUCTIVE REACTANCE. IT, LIKE INDUCTIVE reactance, can be represented as a ray. The capacitive-reactance ray goes in a negative direction and is assigned negative ohmic values. When the capacitive-reactance and inductive-reactance rays are joined at their endpoints (both of which correspond to a reactance of zero), a complete number line is the result, as shown in Fig. 14-1. This line depicts all possible values of reactance.

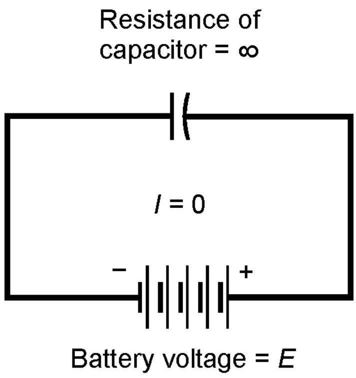
Capacitors and Direct Current

Suppose you have two big, flat metal plates, both of which are excellent electrical conductors. Imagine that you stack them one on top of the other, with only air in between. If you connect a source of dc across the plates (Fig. 14-2), the plates will become electrically charged, and will reach a potential difference equal to the dc source voltage. It won't matter how big or small the plates are; their mutual voltage will always be the same as that of the source, although, if the plates are huge, it will take awhile for them to become fully charged. Once the plates are fully charged, the current will drop to zero.

If you put some insulating material, such as glass, between the plates, their mutual voltage will not change, although the charging time will increase. If you increase the source voltage, the poten-



14-1 Inductive and capacitive reactance can be represented as numerical values (corresponding to ohms multiplied by *j*) along a number line.



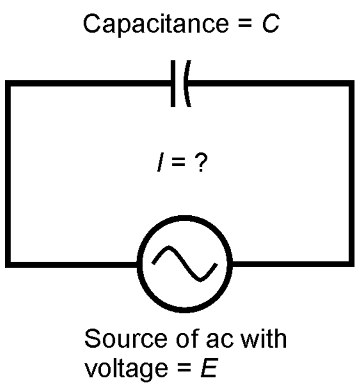
14-2 A capacitor connected across a source of dc.

tial difference between the plates will follow along, more or less rapidly, depending on how large the plates are and on what is between them. If the voltage is increased without limit, arcing will eventually take place. That is, sparks will begin to jump between the plates.

Capacitors and Alternating Current

Now, imagine that the voltage source connected across the plates is changed from dc to ac (Fig. 14-3). Imagine that you can adjust the frequency of this ac from a low value of a few hertz, to hundreds of hertz, to many kilohertz, megahertz, and gigahertz.

At first, the voltage between the plates will follow just about exactly along as the ac source polarity reverses. But the set of plates has a certain amount of capacitance. Perhaps they can charge up fast, if they are small and if the space between them is large, but they can't charge instantaneously. As you increase the frequency of the ac voltage source, there will come a point at which the plates do not get charged up very much before the source polarity reverses. The charge won't have time to get established with each ac cycle. At high ac frequencies, the voltage between the plates will have trouble following the current that is charging and discharging them. Just as the plates begin to get a good charge, the ac current will pass its peak and start to discharge them, pulling electrons out of the negative plate and pumping electrons into the positive plate.



14-3 A capacitor connected across a source of ac.

As the frequency is raised without limit, the set of plates starts to act more and more like a short circuit. When the frequency is low, there is a small charging current, but this quickly drops to zero as the plates become fully charged. As the frequency becomes high, the current flows for more and more of every cycle before dropping off; the charging time remains constant while the *period* of the charging/discharging wave is getting shorter. Eventually, if you keep on increasing the frequency, the period of the wave will be much shorter than the charging/discharging time, and current will flow in and out of the plates in just about the same way as it would flow if the plates were shorted out.

The opposition that the set of plates offers to ac is the *capacitive reactance.* It is measured in ohms, just like inductive reactance, and just like resistance. But it is, by convention, assigned negative values rather than positive ones. Capacitive reactance, denoted X_C, can vary, just as resistance and inductive reactance do, from near zero (when the plates are huge and close together, and/or the frequency is very high) to a few negative ohms, to many negative kilohms or megohms.

Capacitive reactance, like inductive reactance, varies with frequency. But X_C gets larger (negatively) as the frequency goes down. This is the opposite of what happens with inductive reactance, which gets larger (positively) as the frequency goes up.

Often, capacitive reactance is talked about in terms of its *absolute value*, with the minus sign removed. Then we say that the absolute value of X_C increases as the frequency goes down, or that the absolute value of X_C is decreases as the frequency goes up.

Capacitive Reactance and Frequency

In one sense, capacitive reactance behaves like a reflection of inductive reactance. But looked at another way, X_C is an extension of X_L into negative values.

If the frequency of an ac source (in hertz) is given as f, and the capacitance (in farads) is given as C, then the capacitive reactance in ohms, X_C, is calculated as follows:

$$X_C = -1/(2\pi fC)$$

Again, we meet our friend π! And again, for most practical purposes, we can take 2π to be equal to 6.28. Thus, the preceding formula can be expressed like this:

$$X_C = -1/(6.28fC)$$

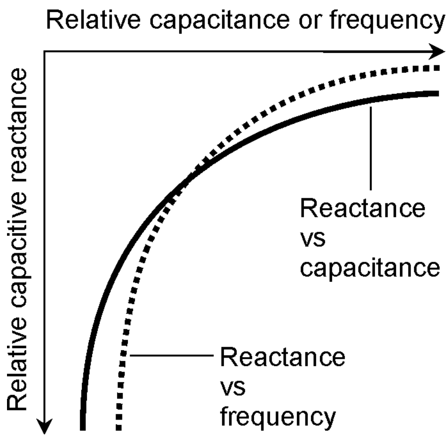
This same formula applies if the frequency, f, is in megahertz and the capacitance, C, is in microfarads.

Capacitive reactance varies *inversely* with the frequency. This means that the function X_C versus f appears as a curve when graphed, and this curve "blows up" as the frequency gets close to zero. Capacitive reactance also varies inversely with the actual value of capacitance, given a fixed frequency. Therefore, the function of X_C versus C also appears as a curve that blows up as the capacitance approaches zero.

The negative of X_C is *inversely proportional* to frequency, and also to capacitance. Relative graphs of these functions are shown in Fig. 14-4.

Problem 14-1

Suppose a capacitor has a value of 0.00100 μF at a frequency of 1.00 MHz. What is the capacitive reactance?



14-4 Capacitive reactance is negatively, and inversely, proportional to capacitance. Capacitive reactance is also negatively, and inversely, proportional to frequency.

Use the formula and plug in the numbers. You can do this directly, because the data is specified in microfarads (millionths) and in megahertz (millions):

$$X_C = -1/(6.28 \times 1.0 \times 0.00100) = -1/(0.00628) = -159\ \Omega$$

This is rounded to three significant figures, because all the data is given to that many digits.

Problem 14-2

What is the capacitive reactance of the preceding capacitor if the frequency decreases to zero (that is, if the voltage source is pure dc)?

In this case, if you plug the numbers into the formula, you get a zero denominator. Mathematicians will tell you that such a quantity is undefined. But we can say that the reactance is negative infinity for all practical purposes.

Problem 14-3

Suppose a capacitor has a reactance of $-100\ \Omega$ at a frequency of 10.0 MHz. What is its capacitance?

In this problem, you need to put the numbers in the formula and solve for the unknown C. Begin with this equation:

$$-100 = -1/(6.28 \times 10.0 \times C)$$

Dividing through by -100, you get:

$$1 = 1/(628 \times 10.0 \times C)$$

Multiply each side of this by C, and you obtain $C = 1/(628 \times 10.0)$. This can be worked out with a calculator. You should find that $C = 0.000159$ to three significant figures. Because the frequency is given in megahertz, the capacitance comes out in microfarads. That means $C = 0.000159\ \mu$F. You can also say it is 159 pF. (Remember that 1 pF = 0.000001 μF.)

Points in the *RC* Plane

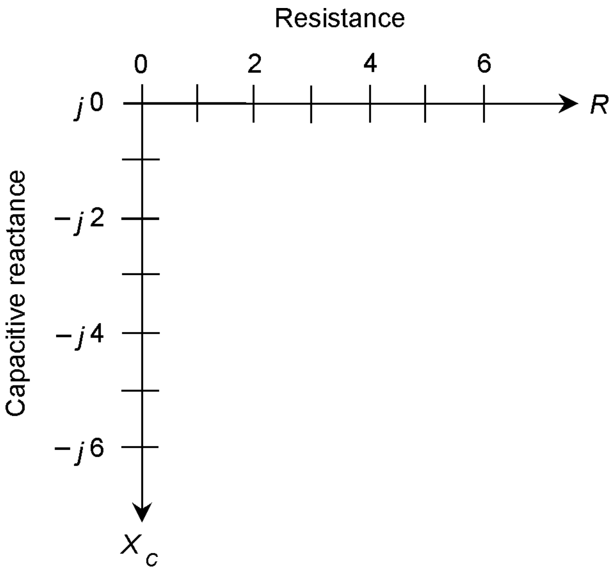
In a circuit containing resistance and capacitive reactance, the characteristics are two-dimensional in a way that is analogous to the situation with the *RL* plane from the previous chapter. The resistance ray and the capacitive-reactance ray can be placed end to end at right angles to make a quarter plane called the *RC plane* (Fig. 14-5). Resistance is plotted horizontally, with increasing values toward the right. Capacitive reactance is plotted downward, with increasingly negative values as you go down.

The combinations of R and X_C in this *RC* plane form impedances. You'll learn about impedance in greater detail in the next chapter. Each point on the *RC* plane corresponds to one and only one impedance. Conversely, each specific impedance coincides with one and only one point on the plane.

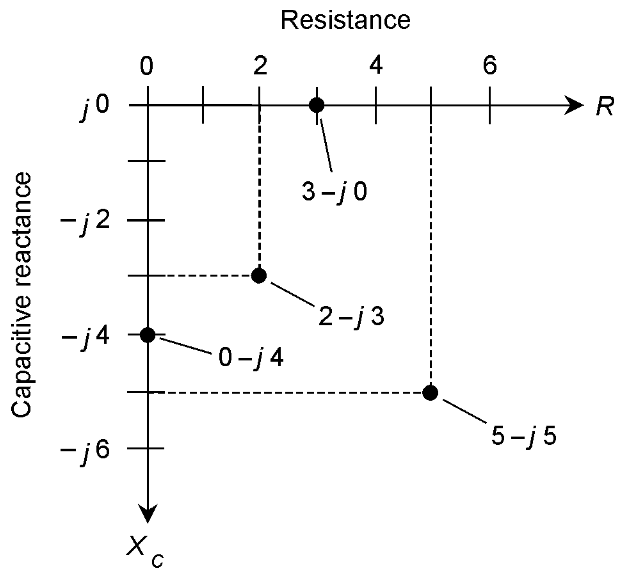
Any impedance that consists of a resistance R and a capacitive reactance X_C can be written in the form $R + jX_C$. Remember that X_C is always negative or zero. Because of this, engineers will often write $R - jX_C$ instead.

If an impedance is a pure resistance R with no reactance, then the complex impedance is $R - j0$ (or $R + j0$; it doesn't matter if j is multiplied by 0!). If $R = 3\ \Omega$ with no reactance, you get an impedance of $3 - j0$, which corresponds to the point $(3,j0)$ on the *RC* plane. If you have a pure capacitive reactance, say $X_C = -4\ \Omega$, then the complex impedance is $0 - j4$, and this is at the point $(0,-j4)$ on the *RC* plane. Again, it's important, for completeness, to write the "0" and not just the "$-j4$." The points for $3 - j0$ and $0 - j4$, and two others, are plotted on the *RC* plane in Fig. 14-6.

In practical circuits, all capacitors have some *leakage resistance.* If the frequency goes to zero (pure dc), a tiny current always flows, because no capacitor has a perfect insulator between its plates. In addition to this, all resistors have a little capacitive reactance because they occupy a finite physical space. So there is no such thing as a mathematically perfect resistor, either. The points $3 - j0$ and



14-5 The quarter plane for capacitive reactance (X_C) and resistance (R). This is also known as the *RC* quarter-plane, or simply as the *RC* plane.



14-6 Four points in the *RC* plane.

$0 - j4$ represent an *ideal resistor* and an *ideal capacitor*, respectively—components that can be worked with in theory, but that you will never see in the real world.

Sometimes, resistance and capacitive reactance are both placed in a circuit deliberately. Then you get impedances such as $2 - j3$ and $5 - j5$, both shown in Fig. 14-6.

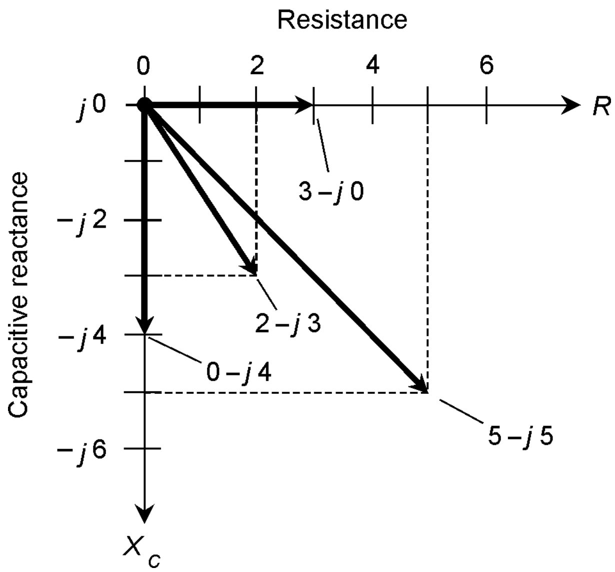
Remember that the values for X_C are reactances, not the actual capacitances. If you raise or lower the frequency, the value of X_C will change. A higher frequency causes X_C to get smaller negatively (closer to zero). A lower frequency causes X_C to get larger negatively (farther from zero, or lower down on the *RC* plane). If the frequency goes to zero, then the capacitive reactance drops off the bottom of the *RC* plane to negative infinity!

Vectors in the *RC* Plane

Recall from the last chapter that *RL* impedances can be represented as vectors. The same is true for *RC* impedances.

In Fig. 14-6, four different complex impedance points are shown. Each point is represented by a certain distance to the right of the origin $(0,j0)$, and a certain displacement downward. The first of these is the resistance, *R*, and the second is the capacitive reactance, X_C. The complex *RC* impedance is a two-dimensional quantity.

Impedance points in the *RC* plane can be rendered as vectors, just as they can in the *RL* plane. Then the points become rays, each with a certain length and direction. The magnitude and direction for a vector, and the coordinates for the point, both uniquely define the same complex impedance. The length of the vector is the distance of the point from the origin, and the direction is the angle measured *clockwise* from the resistance (*R*) line, and specified in *negative degrees.* The equivalent vectors, for the points in Fig. 14-6, are shown in Fig. 14-7.



14-7 Four vectors in the *RC* plane, corresponding to the points shown in Fig. 14-6.

Current Leads Voltage

When ac is driven through a capacitor and starts to increase (in either direction), it takes a fraction of a cycle for the voltage between the plates to follow. Once the current starts decreasing from its maximum peak (in either direction) in the cycle, it again takes a fraction of a cycle for the voltage to follow. The instantaneous voltage can't quite keep up with the instantaneous current, as it does in a pure resistance. Thus, in a circuit containing capacitive reactance, the voltage lags the current in phase. Another, and more often used, way of saying this is that the current *leads* the voltage.

Pure Capacitance

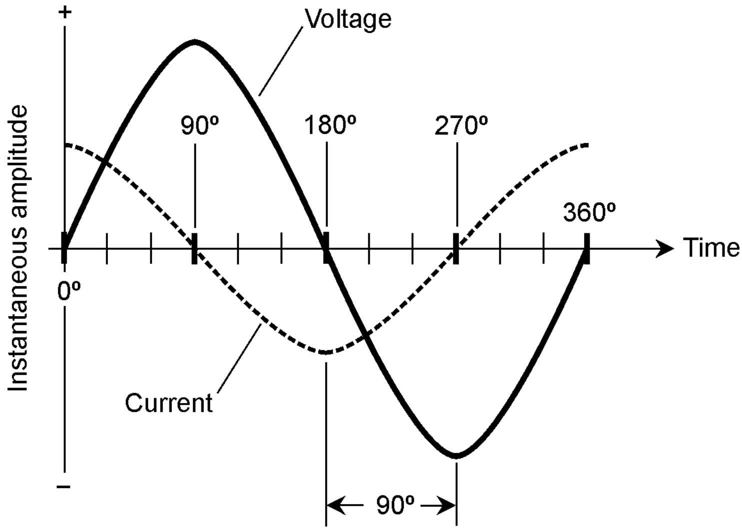
Suppose an ac voltage source is connected across a capacitor. Imagine that the frequency is low enough, and/or the capacitance is small enough, so the absolute value of the capacitive reactance, X_C, is extremely large compared with the resistance, R. Then the current leads the voltage by just about 90° (Fig. 14-8).

The situation depicted in Fig. 14-8 represents a pure capacitive reactance. The vector in the *RC* plane in this situation points straight down. Its angle is −90° from the R axis.

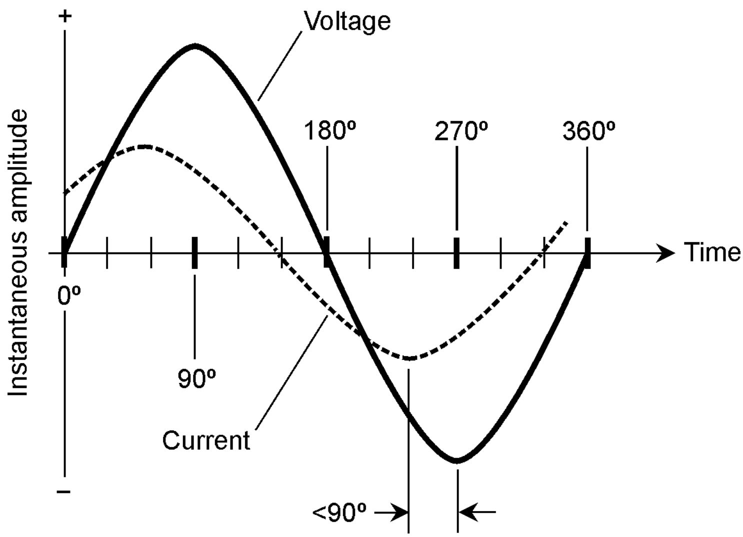
Capacitance and Resistance

When the resistance in a resistance-capacitance circuit is significant compared with the absolute value of the capacitive reactance, the current leads the voltage by something less than 90° (Fig. 14-9). If R is small compared with the absolute value of X_C, the difference is almost a quarter of a cycle. As R gets larger, or as the absolute value of X_C becomes smaller, the phase difference decreases. A circuit containing resistance and capacitance is called an *RC circuit.*

The value of R in an *RC* circuit might increase relative to the absolute value of X_C because resistance is deliberately put into a circuit. It can also happen if the frequency becomes so high that the absolute value of the capacitive reactance drops to a value comparable with the loss resistance in



14-8 In a pure capacitive reactance, the current leads the voltage by 90°.

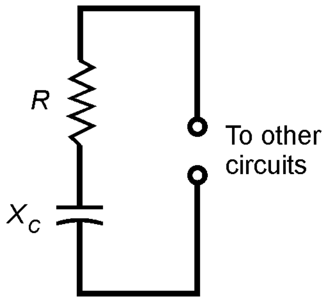


14-9 In a circuit with capacitive reactance and resistance, the current leads the voltage by less than 90°.

the circuit conductors. In either case, the situation can be represented by a resistance, R, in series with a capacitive reactance, X_C (Fig. 14-10).

If you know the values of X_C and R, you can find the *angle of lead*, also called the *RC phase angle*, by plotting the point $R - jX_C$ on the RC plane, drawing the vector from the origin $0 - j0$ out to that point, and then measuring the angle of the vector clockwise from the R axis. You can use a protractor to measure this angle, as you did in the previous chapter for RL phase angles. Or you can use trigonometry to calculate the angle.

As with RL circuits, you need only know the ratio of X_C to R to determine the phase angle. For example, if $X_C = -4\ \Omega$ and $R = 7\ \Omega$, you'll get the same angle as with $X_C = -400\ \Omega$ and $R = 700\ \Omega$, or with $X_C = -16\ \Omega$ and $R = 28\ \Omega$. The phase angle will be the same whenever the ratio of X_C to R is equal to −4:7.



14-10 Schematic representation of a circuit containing resistance and capacitive reactance.

Pure Resistance

As the resistance in an *RC* circuit gets large compared with the absolute value of the capacitive reactance, the angle of lead becomes smaller. The same thing happens if the absolute value of X_C gets small compared with the value of *R*.

When *R* is many times larger than the absolute value of X_C, whatever their actual values, the vector in the *RC* plane points almost along the *R* axis. Then the *RC* phase angle is close to 0°. The voltage comes nearly into phase with the current. The plates of the capacitor do not come anywhere near getting fully charged with each cycle. The capacitor is said to "pass the ac" with very little loss, as if it were shorted out. But it will still have an extremely high X_C for any ac signals at much lower frequencies that might exist across it at the same time. (This property of capacitors can be put to use in electronic circuits. An example is when an engineer wants to let radio-frequency signals get through while blocking signals at audio frequencies.)

Ultimately, if the absolute value of the capacitive reactance gets small enough, the circuit acts as a pure resistance, and the current is in phase with the voltage.

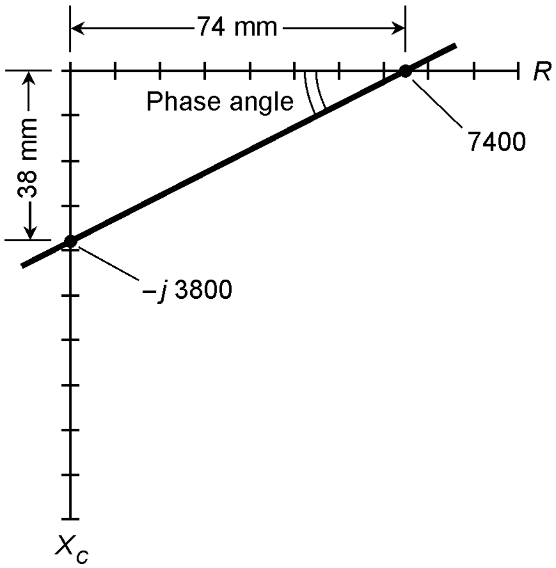
How Much Lead?

If you know the ratio of capacitive reactance to resistance, or X_C/R, in an *RC* circuit, then you can find the phase angle. Of course, you can find this angle if you know the precise values, too.

Pictorial Method

You can use a protractor and a ruler to find phase angles for *RC* circuits, just as you did with *RL* circuits in the previous chapter, as long as the angles aren't too close to 0° or 90°. First, draw a line somewhat longer than 10 cm, going from left to right on the paper. Then, use the protractor to construct a line going somewhat more than 10 cm vertically downward, starting at the left end of the horizontal line. The horizontal line is the *R* axis of an *RC* plane. The line going down is the X_C axis.

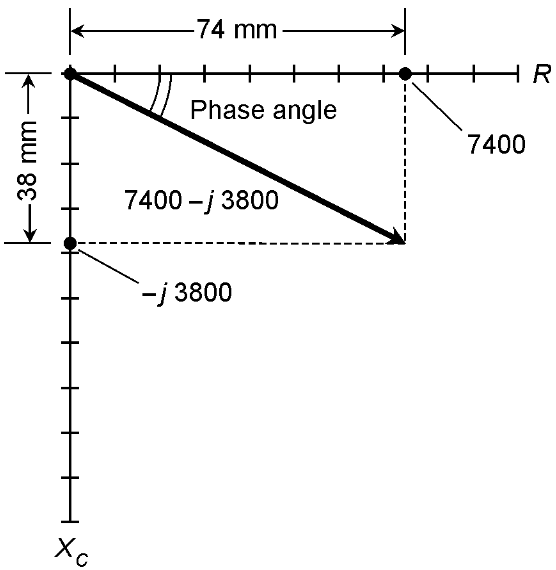
If you know the actual values of X_C and *R*, divide or multiply them by a constant, chosen to make both values fall between −100 and 100. For example, if $X_C = -3800\ \Omega$ and $R = 7400\ \Omega$, divide them both by 100, getting −38 and 74. Plot these points on the lines. The X_C point goes 38 mm down from the intersection point between your two axes. The *R* point goes 74 mm to the right of the intersection point. Next, draw a line connecting the two points, as shown in Fig. 14-11. This line will be at a slant and will form a triangle along with the two axes. This is a right triangle, with the right angle at the origin of the *RC* plane. Measure the angle between the slanted line and the *R* axis. Use the protractor for this. Extend the lines, if necessary, using the ruler, to get a good reading



14-11 Pictorial method of finding phase angle in a circuit containing resistance and capacitive reactance.

on the protractor. This angle will be between 0 and 90°. Multiply this reading by −1 to get the *RC* phase angle. That is, if the protractor shows 27°, the *RC* phase angle is −27°.

The actual vector is found by constructing a rectangle using the origin and your two points, making new perpendicular lines to complete the figure. The vector is the diagonal of this rectangle, running out from the origin (Fig. 14-12). The phase angle is the angle between the *R* axis and this vector, multiplied by −1. It will have the same measure as the angle of the slanted line you constructed in Fig. 14-11.



14-12 Another pictorial method of finding phase angle in a circuit containing resistance and capacitive reactance. This method shows the actual impedance vector.

Trigonometric Method

Using trigonometry, you can determine the *RC* phase angle more precisely than the pictorial method allows. Given the values of X_C and R, the *RC* phase angle is the arctangent of their ratio. Phase angle in *RC* circuits is symbolized by the lowercase Greek letter ϕ, just as it is in *RL* circuits. Here are the formulas:

$$\phi = \tan^{-1}(X_C/R) \qquad \text{or} \qquad \phi = \arctan(X_C/R)$$

When doing problems of this kind, remember to use the *capacitive reactance* values for X_C, and not the capacitance values. This means that, if you are given the capacitance, you must use the formula for X_C in terms of capacitance and frequency and then calculate the phase angle. You should get angles that come out negative or zero. This indicates that they're *RC* phase angles rather than *RL* phase angles (which are always positive or zero).

Problem 14-4

Suppose the capacitive reactance in an *RC* circuit is −3800 Ω and the resistance is 7400 Ω. What is the phase angle?

Find the ratio $X_C/R = -3800/7400$. The calculator display should show you something like −0.513513513. Find the arctangent, or $\tan^{-1}$, getting a phase angle of −27.18111109° on the calculator display. Round this off to −27.18°.

Problem 14-5

Suppose an *RC* circuit works at a frequency of 3.50 MHz. It has a resistance of 130 Ω and a capacitance of 150 pF. What is the phase angle?

First, find the capacitive reactance for a capacitor of 150 pF at 3.50 MHz. Convert the capacitance to microfarads, getting $C = 0.000150\ \mu\text{F}$. Remember that *micro*farads go with *mega*hertz (millionths go with millions to cancel each other out). Then:

$$\begin{aligned} X_C &= -1/(6.28 \times 3.50 \times 0.000150) \\ &= -1/0.003297 = -303\ \Omega \end{aligned}$$

Now you can find the ratio $X_C/R = -303/130 = -2.33$. The phase angle is equal to the arctangent of −2.33, or −66.8°.

Problem 14-6

What is the phase angle in the preceding circuit if the frequency is raised to 7.10 MHz?

You need to find the new value for X_C, because it will change as a result of the frequency change. Calculating:

$$\begin{aligned} X_C &= -1/(6.28 \times 7.10 \times 0.000150) \\ &= -1/0.006688 = -150\ \Omega \end{aligned}$$

The ratio X_C/R in this case is equal to −150/130, or −1.15. The phase angle is the arctangent of −1.15, which turns out to be −49.0°.

Quiz

Refer to the text in this chapter if necessary. A good score is at least 18 correct. Answers are in the back of the book.

1. As the size of the plates in a capacitor increases, all other things being equal,
   (a) the value of X_C increases negatively.
   (b) the value of X_C decreases negatively.
   (c) the value of X_C does not change.
   (d) we cannot say what happens to X_C without more data.

2. If the dielectric material between the plates of a capacitor is changed, all other things being equal,
   (a) the value of X_C increases negatively.
   (b) the value of X_C decreases negatively.
   (c) the value of X_C does not change.
   (d) we cannot say what happens to X_C without more data.

3. As the frequency of a wave gets lower, all other things being equal, the value of X_C for a capacitor
   (a) increases negatively.
   (b) decreases negatively.
   (c) does not change.
   (d) depends on the current.

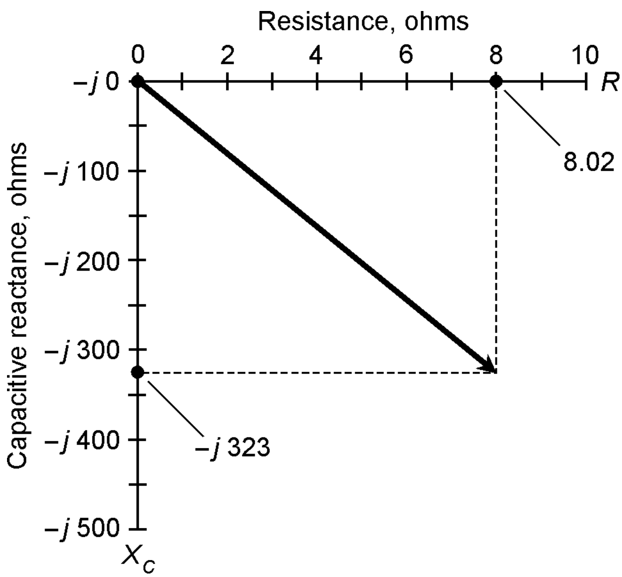
4. What is the reactance of a 330-pF capacitor at 800 kHz?
   (a) −1.66 Ω
   (b) −0.00166 Ω
   (c) −603 Ω
   (d) −603 kΩ

5. Suppose a capacitor has a reactance of −4.50 Ω at 377 Hz. What is its capacitance?
   (a) 9.39 µF
   (b) 93.9 µF
   (c) 7.42 µF
   (d) 74.2 µF

6. Suppose a 47-µF capacitor has a reactance of −47 Ω. What is the frequency?
   (a) 72 Hz
   (b) 7.2 MHz
   (c) 0.000072 Hz
   (d) 7.2 Hz

7. Suppose a capacitor has $X_C = -8800\ \Omega$ at $f = 830$ kHz. What is C?
   (a) 2.18 μF
   (b) 21.8 pF
   (c) 0.00218 μF
   (d) 2.18 pF
8. Suppose a capacitor has $C = 166$ pF at $f = 400$ kHz. What is X_C?
   (a) −2.4 kΩ
   (b) −2.4 Ω
   (c) $-2.4 \times 10^{-6}\ \Omega$
   (d) −2.4 MΩ
9. Suppose a capacitor has $C = 4700$ μF and $X_C = -33\ \Omega$. What is f?
   (a) 1.0 Hz
   (b) 10 Hz
   (c) 1.0 kHz
   (d) 10 kHz
10. Each point in the *RC* plane
   (a) corresponds to a unique inductance.
   (b) corresponds to a unique capacitance.
   (c) corresponds to a unique combination of resistance and capacitance.
   (d) corresponds to a unique combination of resistance and reactance.
11. If R increases in an *RC* circuit, but X_C is always zero, the vector in the *RC* plane will
   (a) rotate clockwise.
   (b) rotate counterclockwise.
   (c) always point straight toward the right.
   (d) always point straight down.
12. If the resistance R increases in an *RC* circuit, but the capacitance and the frequency are nonzero and constant, then the vector in the *RC* plane will
   (a) get longer and rotate clockwise.
   (b) get longer and rotate counterclockwise.
   (c) get shorter and rotate clockwise.
   (d) get shorter and rotate counterclockwise.
13. Each complex impedance value $R - jX_C$
   (a) represents a unique combination of resistance and capacitance.
   (b) represents a unique combination of resistance and reactance.
   (c) represents a unique combination of resistance and frequency.
   (d) All of the above are true.

14. In an *RC* circuit, as the ratio X_C/R approaches zero, the phase angle
    (a) approaches −90°.
    (b) approaches 0°.
    (c) stays the same.
    (d) cannot be found.
15. In a purely resistive circuit, the phase angle is
    (a) increasing.
    (b) decreasing.
    (c) 0°.
    (d) −90°.
16. If $X_C/R = -1$, then what is the phase angle?
    (a) 0°
    (b) −45°
    (c) −90°
    (d) Impossible to find because there's not enough data given
17. In Fig. 14-13, the impedance shown is
    (a) $8.02 + j323$.
    (b) $323 + j8.02$.
    (c) $8.02 - j323$.
    (d) $323 - j8.02$.



14-13 Illustration for Quiz Questions 17 and 18.

18. In Fig. 14-13, note that the R and X_C scale divisions are not the same size. What is the actual phase angle?
    (a) −1.42°
    (b) About −60°, from the looks of it
    (c) −58.9°
    (d) −88.6°

19. Suppose an RC circuit consists of a 150-pF capacitor and a 330-Ω resistor in series. What is the phase angle at a frequency of 1.34 MHz?
    (a) −67.4°
    (b) −22.6°
    (c) −24.4°
    (d) −65.6°

20. Suppose an RC circuit has a capacitance of 0.015 μF. The resistance is 52 Ω. What is the phase angle at 90 kHz?
    (a) −24°
    (b) −0.017°
    (c) −66°
    (d) None of the above

15 CHAPTER

Impedance and Admittance

IN THIS CHAPTER, A COMPLETE, WORKING DEFINITION OF COMPLEX IMPEDANCE IS DEVELOPED. YOU'LL also get acquainted with *admittance*, the extent to which an ac circuit allows (or admits) current flow, rather than impeding it. As we develop these concepts, let's review, and then expand on, some of the material presented in the previous couple of chapters.

Imaginary Numbers

Have you been wondering what j actually means in expressions of impedance? Well, j is nothing but a number: the positive square root of −1. There's a negative square root of −1, too, and it is equal to $-j$. When either j or $-j$ is multiplied by itself, the result is −1. (Pure mathematicians often denote these same numbers as i or $-i$.)

The positive square root of −1 is known as the *unit imaginary number.* The *set of imaginary numbers* is composed of real-number multiples of j or $-j$. Some examples are $j4$, $j35.79$, $-j25.76$, and $-j25{,}000$.

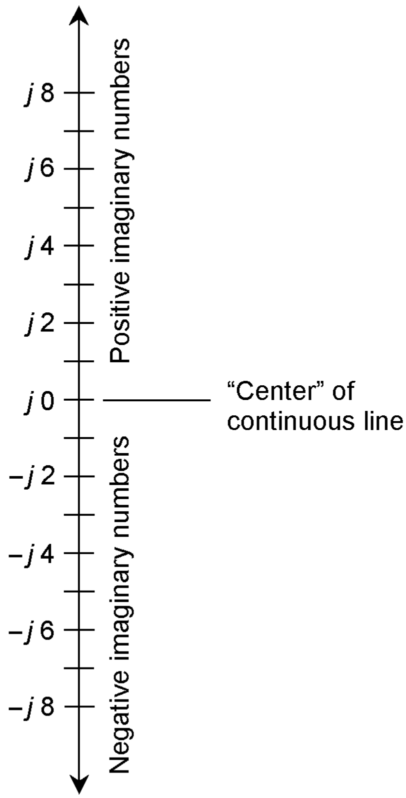
The square of an imaginary number is always negative. Some people have trouble grasping this, but when you think long and hard about it, all numbers are abstractions. Imaginary numbers are no more imaginary (and no less real) than so-called real numbers such as 4, 35.79, −25.76, or −25,000.

The unit imaginary number j can be multiplied by any real number on a conventional *real number line.* If you do this for all the real numbers on the real number line, you get an *imaginary number line* (Fig. 15-1). The imaginary number line should be oriented at a right angle to the real number line when you want to graphically portray real and imaginary numbers at the same time.

In electronics, real numbers represent resistances. Imaginary numbers represent reactances.

Complex Numbers

When you add a real number and an imaginary number, you get a *complex number.* In this context, the term *complex* does not mean "complicated." A better word would be *composite.* Examples are



15-1 The imaginary number line.

$4 + j5$, $8 - j7$, $-7 + j13$, and $-6 - j87$. The set of complex numbers needs two dimensions—a *plane*—to be graphically defined.

Adding and Subtracting Complex Numbers

Adding complex numbers is just a matter of adding the real parts and the complex parts separately. For example, the sum of $4 + j7$ and $45 - j83$ works out like this:

$$\begin{aligned} &(4 + 45) + j(7 - 83) \\ &= 49 + j(-76) \\ &= 49 - j76 \end{aligned}$$

Subtracting complex numbers is a little more involved; it's best to convert a difference to a sum. For example, the difference $(4 + j7) - (45 - j83)$ can be found by multiplying the second complex number by -1 and then adding the result:

$$\begin{aligned} &(4 + j7) - (45 - j83) \\ &= (4 + j7) + [-1(45 - j83)] \\ &= (4 + j7) + (-45 + j83) \\ &= -41 + j90 \end{aligned}$$

Multiplying Complex Numbers

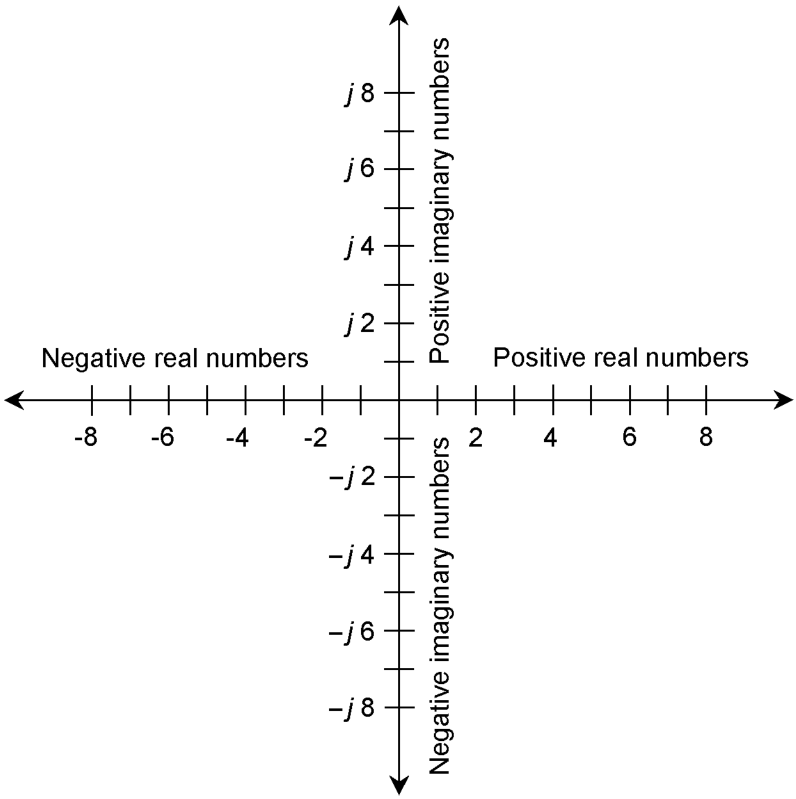
When you multiply these numbers, you should treat them as sums of number pairs, that is, as *binomials.* It's easier to give the general formula than to work with specifics here. If a, b, c, and d are real numbers (positive, negative, or zero), then:

$$
\begin{aligned}
&(a + jb)\,(c + jd)\\
&= ac + jad + jbc + j^2bd\\
&= (ac - bd) + j(ad + bc)
\end{aligned}
$$

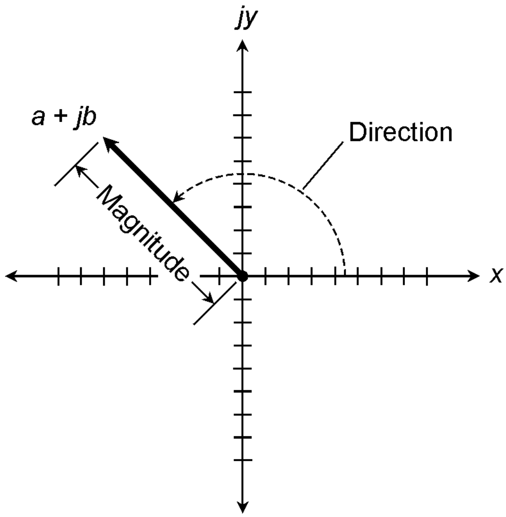
Fortunately, you won't encounter complex number multiplication problems very often in electronics. Nevertheless, a working knowledge of how complex numbers multiply can help you get a solid grasp of them.

The Complex Number Plane

A complete *complex number plane* is made by taking the real and imaginary number lines and placing them together, at right angles, so that they intersect at the zero points, 0 and $j0$. This is shown in Fig. 15-2. The result is a *Cartesian coordinate plane*, just like the ones people use to make graphs of everyday things such as stock price versus time.



15-2 The complex number plane.



15-3 Magnitude and direction of a vector in the complex number plane.

Complex Number Vectors

Complex numbers can also be represented as vectors. This gives each complex number a unique *magnitude* and a unique *direction.* The magnitude is the distance of the point $a + jb$ from the origin $0 + j0$. The direction is the angle of the vector, expressed counterclockwise from the positive real-number axis. This is shown in Fig. 15-3.

Absolute Value

The *absolute value* of a complex number $a + jb$ is the length, or magnitude, of its vector in the complex plane, measured from the origin (0,0) to the point (a,b).

In the case of a *pure real number* $a + j0$, the absolute value is simply the real number itself, a, if a is positive. If a is negative, then the absolute value of $a + j0$ is equal to $-a$.

In the case of a *pure imaginary number* $0 + jb$, the absolute value is equal to b, if b (a real number) is positive. If b is negative, the absolute value of $0 + jb$ is equal to $-b$.

If the number $a + jb$ is neither pure real or pure imaginary, the absolute value must be found by using a formula. First, square both a and b. Then add them. Finally, take the square root. This is the length, c, of the vector $a + jb$. The situation is illustrated in Fig. 15-4.

Problem 15-1

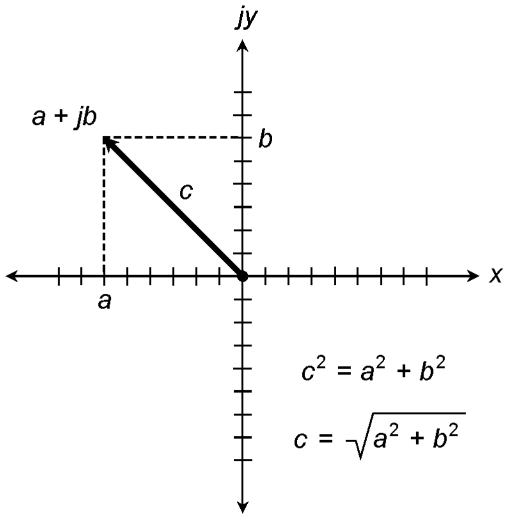
Find the absolute value of the complex number $-22 - j0$.

This is a pure real number. Actually, it is the same as $-22 + j0$, because $j0 = 0$. Therefore, the absolute value of this complex number is $-(-22) = 22$.

Problem 15-2

Find the absolute value of $0 - j34$.

This is a pure imaginary number. The value of b in this case is -34, because $0 - j34 = 0 + j(-34)$. Therefore, the absolute value is $-(-34) = 34$.



15-4 Calculation of absolute value, or vector length. Here, the vector length is represented by *c*.

Problem 15-3

Find the absolute value of $3 - j4$.

In this number, $a = 3$ and $b = -4$. Squaring both of these, and adding the results, gives us $3^2 + (-4)^2 = 9 + 16 = 25$. The square root of 25 is 5. Therefore, the absolute value of this complex number is 5.

The *RX* Plane

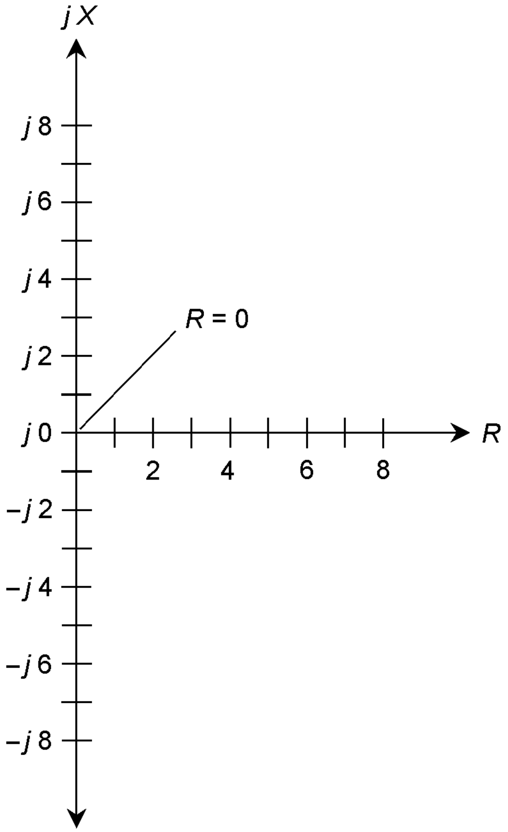
Recall the planes for resistance (R) and inductive reactance (X_L) from Chap. 13. This is the same as the upper-right quadrant of the complex number plane shown in Fig. 15-2. Similarly, the plane for resistance and capacitive reactance (X_C) is the same as the lower-right quadrant of the complex number plane. Resistances are represented by nonnegative real numbers. Reactances, whether they are inductive (positive) or capacitive (negative), correspond to imaginary numbers.

No Negative Resistance

There is no such thing, strictly speaking, as negative resistance. You cannot have anything better than a perfect conductor. In some cases, a supply of direct current, such as a battery, can be treated as a negative resistance; in other cases, you can have a device that acts as if its resistance were negative under certain changing (or dynamic) conditions. But for most practical applications in the *RX plane*, the resistance value is always positive. You can remove the negative axis, along with the upper-left and lower-left quadrants, of the complex number plane, obtaining a half plane, as shown in Fig. 15-5, and still get a complete set of coordinates for depicting complex impedances.

"Negative Inductors" and "Negative Capacitors"

Capacitive reactance, X_C, is effectively an extension of inductive reactance, X_L, into the realm of negatives. Capacitors act like "negative inductors." It's equally true to say that inductors act like "negative capacitors," because the negative of a negative number is a positive number. Reactance can vary from extremely large negative values, through zero, to extremely large positive values.



15-5 The complex impedance plane, also called the resistance-reactance (RX) plane.

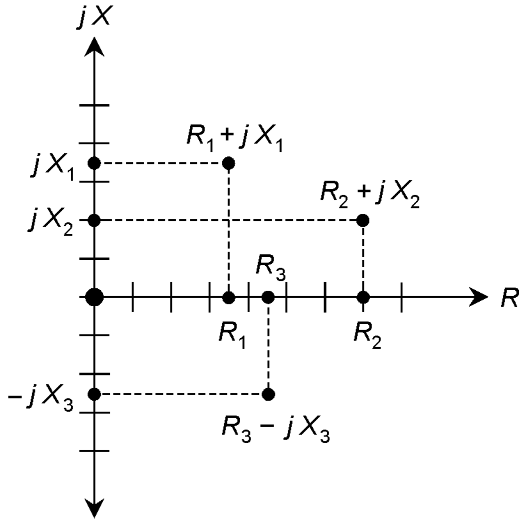
Vector Representation of Impedance

Any impedance $R + jX$ can be represented by a complex number of the form $a + jb$. Just let $R = a$ and $X = b$. Now try to envision how the impedance vector changes as either R or X, or both, are varied. If X remains constant, an increase in R causes the vector to get longer. If R remains constant and X_L gets larger, the vector grows longer. If R stays the same but X_C gets larger negatively, the vector grows longer.

Think of the point $R + jX$ moving around in the RX plane, and imagine where the corresponding points on the axes lie. These points can be found by drawing dashed lines from the point $R + jX$ to the R and X axes, so that the dashed lines intersect the axes at right angles. Some examples are shown in Fig. 15-6.

Now think of the points for R and X moving toward the right and left, or up and down, on their axes. Imagine what happens to the point $R + jX$ in various scenarios. This is how impedance changes as the resistance and reactance in a circuit are varied.

Resistance is one-dimensional. Reactance is also one-dimensional. But impedance is two-dimensional. To fully define impedance, you must render it on a two-dimensional coordinate system such as the RX plane. The resistance and the reactance can change independently of one another.



15-6 Some points in the complex impedance plane, and their resistive and reactive components on the axes.

Absolute-Value Impedance

You'll occasionally read or hear that the "impedance" of some device or component is a certain number of ohms. For example, in audio electronics, there are "8-Ω" speakers and "600-Ω" amplifier inputs. How, you ask, can manufacturers quote a single number for a quantity that is two-dimensional and needs two numbers to be completely expressed?

That's a good question, and there are two answers. First, figures like this refer to devices that have *purely resistive impedances*, also known as *nonreactive impedances.* Thus, the 8-Ω speaker really has a complex impedance of $8 + j0$, and the 600-Ω input circuit is designed to operate with a complex impedance at, or near, $600 + j0$. Second, you can talk about the length of the impedance vector (that is, the absolute value of the complex impedance), calling this a certain number of ohms. If you talk about impedance this way, however, you are being ambiguous. There can exist an infinite number of different vectors of any given length in the *RX* plane.

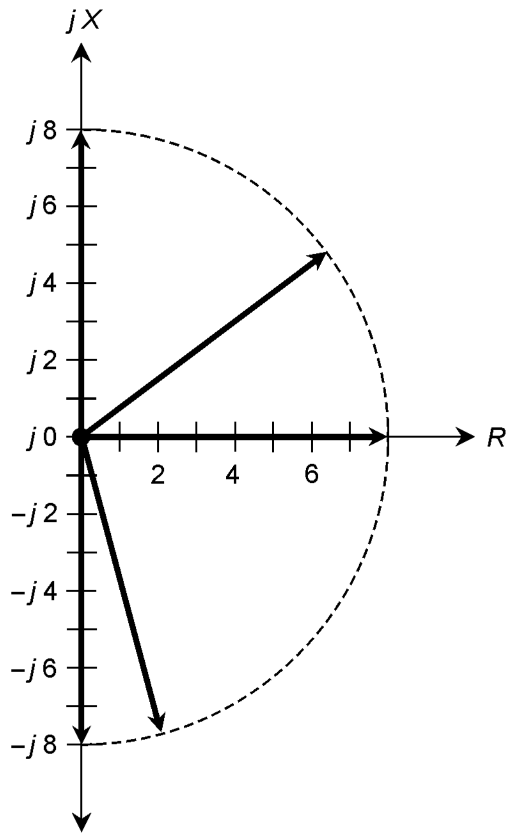
Sometimes, the uppercase italic letter Z is used in place of the word *impedance* in general discussions. This is what engineers mean when they say things like "$Z = 50\ \Omega$" or "$Z = 300\ \Omega$ nonreactive." In this context, if no specific impedance is given, "$Z = 8\ \Omega$" can theoretically refer to $8 + j0$, $0 + j8$, $0 - j8$, or any other complex impedance point on a half circle consisting of all points 8 units from $0 + j0$. This is shown in Fig. 15-7.

Problem 15-4

Name seven different complex impedances that can theoretically be meant by the expression "$Z = 10\ \Omega$."

It's easy name three: $0 + j10$, $10 + j0$, and $0 - j10$. These represent pure inductance, pure resistance, and pure capacitance, respectively.

A right triangle can exist having sides in a ratio of 6:8:10 units. This is true because $6^2 + 8^2 = 10^2$. (Check it and see!) Therefore, you can have $6 + j8$, $6 - j8$, $8 + j6$, and $8 - j6$, all complex impedances whose absolute value is 10.



15-7 Vectors representing an absolute-value impedance of 8 Ω.

Characteristic Impedance

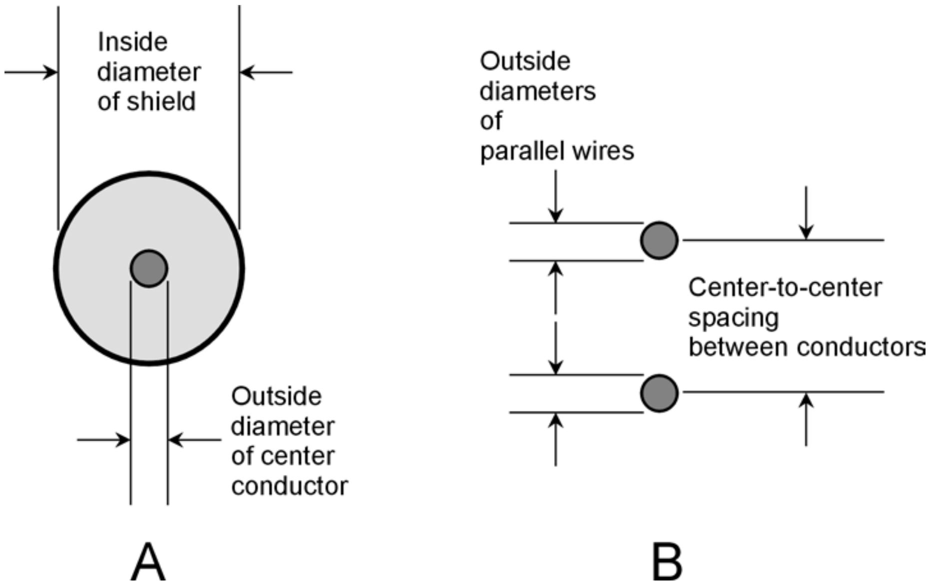
There is a rather exotic property of certain electronic components that you'll sometimes hear or read about. It is called *characteristic impedance* or *surge impedance*, and is symbolized Z_o. It is a specification of an important property of *transmission lines.* It can always be expressed as a positive real number, in ohms.

Transmission Lines

When it is necessary to get energy or signals from one place to another, a transmission line is required. These almost always take either of two forms, *coaxial* or *two-wire* (also called *parallel-wire*). Cross-sectional renditions of both types are shown in Fig. 15-8. Examples of transmission lines include the "ribbon" that goes from a television antenna to the receiver, the cable running from a hi-fi amplifier to the speakers, and the set of wires that carries electricity over the countryside.

Factors Affecting Z_o

The Z_o of a parallel-wire transmission line depends on the diameter of the wires, on the spacing between the wires, and on the nature of the insulating material separating the wires. In general, the Z_o



15-8 Edge-on views of coaxial transmission line (A) and parallel-wire line (B). In either type of line, Z_o depends on the conductor diameters and spacing, and on the nature of the dielectric material between the conductors. See text for discussion.

increases as the wire diameter gets smaller, and decreases as the wire diameter gets larger, all other things being equal.

In a coaxial line, as the center conductor gets thicker, the Z_o decreases if the shield stays the same size. If the center conductor stays the same size and the *shield* increases in diameter, the Z_o increases.

For either type of line, the Z_o increases as the spacing between wires, or between the center conductor and the shield, gets larger. The Z_o decreases as the spacing is reduced. Solid dielectric materials such as polyethylene reduce the Z_o of a transmission line, compared with air or a vacuum, when placed between the conductors.

An Example of Z_o in Practice

In rigorous terms, the ideal characteristic impedance for a transmission line is determined according to the nature of the *load* with which the line works.

For a system having a purely resistive impedance of a certain number of ohms, the best line Z_o value is that same number of ohms. If the load impedance is much different from the characteristic impedance of the transmission line, excessive power is wasted in heating up the transmission line.

Imagine that you have a so-called 300-Ω frequency-modulation (FM) receiving antenna, such as the folded-dipole type that you can mount indoors. Suppose that you want the best possible reception. Of course, you should choose a good location for the antenna. You should make sure that the transmission line between your radio and the antenna is as short as possible. But you should also be sure that you purchase 300-Ω TV ribbon. It has a value of Z_o that has been optimized for use with antennas whose impedances are close to $300 + j0$.

Impedance matching is the process of making sure that the impedance of a load (such as an antenna) is purely resistive, with an ohmic value equal to the characteristic impedance of the transmission line connected to it. This concept will be discussed in more detail in the next chapter.

Conductance

In an ac circuit, electrical *conductance* works the same way as it does in a dc circuit. Conductance is symbolized by the capital letter G. It was introduced in Chap. 2. The relationship between conductance and resistance is simple: $G = 1/R$. The standard unit of conductance is the *siemens.* The larger the value of conductance, the smaller the resistance, and the more current will flow. Conversely, the smaller the value of G, the greater the value of R, and the less current will flow.

Susceptance

Sometimes, you'll come across the term *susceptance* in reference to ac circuits. Susceptance is symbolized by the capital letter B. It is the reciprocal of reactance. Susceptance can be either capacitive or inductive. These quantities are symbolized as B_C and B_L, respectively. Therefore we have these two relations:

$$B_C = 1/X_C$$
$$B_L = 1/X_L$$

All values of B theoretically contain the j operator, just as do all values of X. But when it comes to finding reciprocals of quantities containing j, things get tricky. The reciprocal of j is equal to its negative! Expressed mathematically, we have these two facts:

$$1/j = -j$$
$$1/(-j) = j$$

As a result of these properties of j, the sign reverses whenever you find a susceptance value in terms of a reactance value. When expressed in terms of j, inductive susceptance is negative imaginary, and capacitive susceptance is positive imaginary—just the opposite situation from inductive reactance and capacitive reactance.

Suppose you have an inductive reactance of 2 Ω. This is expressed in imaginary terms as $j2$. To find the inductive susceptance, you must find $1/(j2)$. Mathematically, this expression can be converted to a real-number multiple of j in the following manner:

$$\begin{aligned} 1/(j2) &= (1/j)(1/2) \\ &= (1/j)0.5 \\ &= -j0.5 \end{aligned}$$

Now suppose you have a capacitive reactance of 10 Ω. This is expressed in imaginary terms as $-j10$. To find the capacitive susceptance, you must find $1/(-j10)$. Here's how this can be converted to the straightforward product of j and a real number:

$$\begin{aligned}1/(-j10) &= (1/-j)(^1\!/\!_{10})\\ &= (1/-j)0.1\\ &= j0.1\end{aligned}$$

When you want to find an imaginary value of susceptance in terms of an imaginary value of reactance, first take the reciprocal of the real-number part of the expression, and then multiply the result by −1.

Problem 15-5

Suppose you have a capacitor of 100 pF at a frequency of 3.00 MHz. What is B_C?

First, find X_C by the formula for capacitive reactance:

$$X_C = -1/(6.28fC)$$

Note that 100 pF = 0.000100 μF. Therefore:

$$\begin{aligned}X_C &= -1/(6.28 \times 3.00 \times 0.000100)\\ &= -1/0.001884 = -531\ \Omega\end{aligned}$$

The imaginary value of X_C is equal to $-j531$. The susceptance, B_C, is equal to $1/X_C$. Thus, $B_C = 1/(-j531) = j0.00188$, rounded to three significant figures.

The general formula for capacitive susceptance in siemens, in terms of frequency in hertz and capacitance in farads, is:

$$B_C = 6.28fC$$

This formula also works for frequencies in megahertz and capacitances in microfarads.

Problem 15-6

Suppose an inductor has $L = 163$ μH at a frequency of 887 kHz. What is B_L?

Note that 887 kHz = 0.887 MHz. You can calculate X_L from the formula for inductive reactance:

$$\begin{aligned}X_L &= 6.28fL\\ &= 6.28 \times 0.887 \times 163\\ &= 908\ \Omega\end{aligned}$$

The imaginary value of X_L is equal to $j908$. The susceptance, B_L = is equal to $1/X_L$. It follows that $B_L = -1/j908 = -j0.00110$.

The general formula for inductive susceptance in siemens, in terms of frequency in hertz and inductance in henrys, is:

$$B_L = -1/(6.28fL)$$

This formula also works for frequencies in kilohertz and inductances in millihenrys, and for frequencies in megahertz and inductances in microhenrys.

Admittance

Real-number conductance and imaginary-number susceptance combine to form *complex admittance*, symbolized by the capital letter Y. This is a complete expression of the extent to which a circuit allows ac to flow.

As the absolute value of complex impedance gets larger, the absolute value of complex admittance becomes smaller, in general. Huge impedances correspond to tiny admittances, and vice versa.

Admittances are written in complex form just like impedances. But you need to keep track of which quantity you're talking about! This will be obvious if you use the symbol, such as $Y = 3 - j0.5$ or $Y = 7 + j3$. When you see Y instead of Z, you know that negative j factors (such as in the quantity $3 - j0.5$) mean there is a net inductance in the circuit, and positive j factors (such as in the quantity $7 + j3$) mean there is net capacitance.

Admittance is the complex composite of conductance and susceptance. Thus, complex admittance values always take the form $Y = G + jB$. When the j factor is negative, a complex admittance may appear in the form $Y = G - jB$.

Do you remember how resistances combine with reactances in series to form complex impedances? In Chaps. 13 and 14, you saw series *RL* and *RC* circuits. Did you wonder why parallel circuits were ignored in those discussions? The reason was the fact that admittance, not impedance, is best for working with parallel ac circuits. Resistance and reactance combine in a messy fashion in parallel circuits. But conductance (G) and susceptance (B) merely add together in parallel circuits, yielding admittance (Y). Parallel circuit analysis is covered in detail in the next chapter.

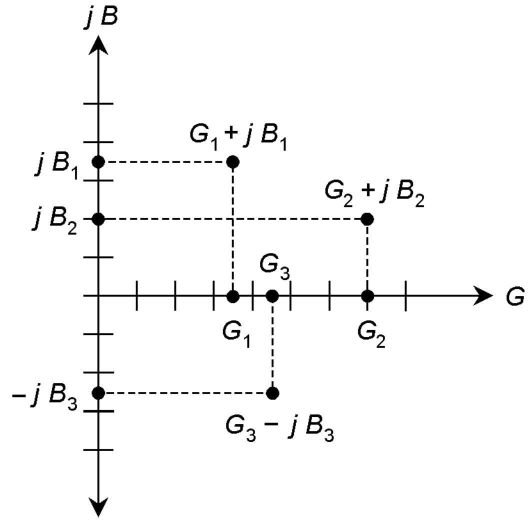
The *GB* Plane

Admittance can be depicted on a plane similar to the complex impedance (*RX*) plane. Actually, it's a half plane, because there is ordinarily no such thing as negative conductance. (You can't have a component that conducts worse than not at all.) Conductance is plotted along the horizontal, or *G*, axis on this coordinate half plane, and susceptance is plotted along the *B* axis. The *GB plane* is shown in Fig. 15-9, with several points plotted.

It's Inside Out

The *GB* plane looks superficially identical to the *RX* plane. But mathematically, the two could not be more different! The *GB* plane is mathematically inside out with respect to the *RX* plane. The center, or origin, of the *GB* plane represents the point at which there is no conduction for dc or for ac. It is the zero-admittance point, rather than the zero-impedance point. In the *RX* plane, the origin represents a perfect short circuit, but in the *GB* plane, the origin corresponds to a perfect open circuit.

As you move out toward the right (east) along the *G*, or conductance, axis of the *GB* plane, the conductance improves, and the current gets greater. When you move upward (north) along the *jB* axis from the origin, you have ever-increasing positive (capacitive) susceptance. When you go down (south) along the *jB* axis from the origin, you encounter increasingly negative (inductive) susceptance.

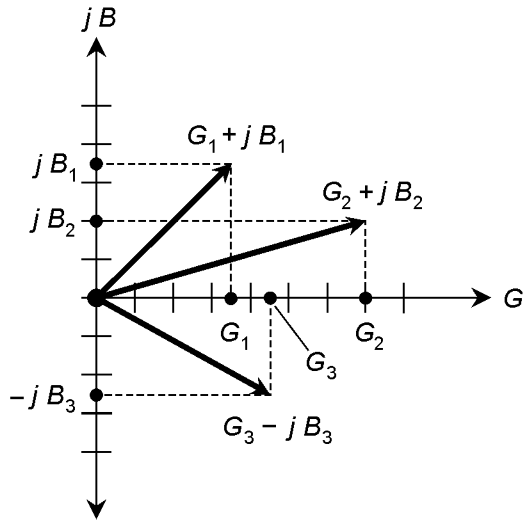


15-9 Some points in the complex admittance plane, and their conductive and susceptive components on the axes.

Vector Representation of Admittance

Complex admittances can be shown as vectors, just as can complex impedances. In Fig. 15-10, the points from Fig. 15-9 are rendered as vectors.

Generally, long vectors in the *GB* plane indicate large currents, and short vectors indicate small currents. Imagine a point moving around on the *GB* plane, and think of the vector getting longer and shorter and changing direction. Vectors pointing generally northeast, or upward and to the right, correspond to conductances and capacitances in parallel. Vectors pointing in a more or less southeasterly direction, or downward and to the right, are conductances and inductances in parallel.



15-10 Vectors representing the points of Fig. 15-9.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 or more correct. Answers are in the back of the book.

1. The square of an imaginary number
   (a) can never be negative.
   (b) can never be positive.
   (c) can be either positive or negative.
   (d) is equal to j.
2. A complex number
   (a) is the same thing as an imaginary number.
   (b) has a real-number part and an imaginary-number part.
   (c) is one-dimensional.
   (d) is a concept reserved for elite mathematicians.
3. What is the sum of $3 + j7$ and $-3 - j7$?
   (a) $0 + j0$
   (b) $6 + j14$
   (c) $-6 - j14$
   (d) $0 - j14$
4. What is $(-5 + j7) - (4 - j5)$?
   (a) $-1 + j2$
   (b) $-9 - j2$
   (c) $-1 - j2$
   (d) $-9 + j12$
5. What is the product $(-4 - j7)(6 - j2)$?
   (a) $24 - j14$
   (b) $-38 - j34$
   (c) $-24 - j14$
   (d) $-24 + j14$
6. What is the magnitude of the vector $18 - j24$?
   (a) 6
   (b) 21
   (c) 30
   (d) 52
7. The complex impedance value $5 + j0$ represents
   (a) a pure resistance.
   (b) a pure inductance.

(c) a pure capacitance.
(d) an inductance combined with a capacitance.

8. The complex impedance value $0 - j22$ represents
(a) a pure resistance.
(b) a pure inductance.
(c) a pure capacitance.
(d) an inductance combined with a resistance.

9. What is the absolute-value impedance of $3.0 - j6.0$?
(a) $Z = 9.0\ \Omega$
(b) $Z = 3.0\ \Omega$
(c) $Z = 45\ \Omega$
(d) $Z = 6.7\ \Omega$

10. What is the absolute-value impedance of $50 - j235$?
(a) $Z = 240\ \Omega$
(b) $Z = 58{,}000\ \Omega$
(c) $Z = 285\ \Omega$
(d) $Z = -185\ \Omega$

11. If the center conductor of a coaxial cable is made to have a smaller diameter, all other things being equal, what will happen to the Z_o of the transmission line?
(a) It will increase.
(b) It will decrease.
(c) It will not change.
(d) There is no way to determine this without knowing the actual dimensions.

12. If a device is said to have an impedance of $Z = 100\ \Omega$, you can reasonably expect that this indicates
(a) $R + jX = 100 + j0$.
(b) $R + jX = 0 + j100$.
(c) $R + jX = 100 + j100$.
(d) the reactance and the resistance add up to $100\ \Omega$.

13. Suppose a capacitor has a value of $0.050\ \mu F$ at 665 kHz. What is the capacitive susceptance, stated as an imaginary number?
(a) $B_C = j4.79$
(b) $B_C = -j4.79$
(c) $B_C = j0.209$
(d) $B_C = -j0.209$

14. An inductor has a value of 44 mH at 60 Hz. What is the inductive susceptance, stated as an imaginary number?
    (a) $B_L = -j0.060$
    (b) $B_L = j0.060$
    (c) $B_L = -j17$
    (d) $B_L = j17$
15. Susceptance and conductance add to form
    (a) complex impedance.
    (b) complex inductance.
    (c) complex reactance.
    (d) complex admittance.
16. Absolute-value impedance is equal to the square root of which of the following?
    (a) $G^2 + B^2$
    (b) $R^2 + X^2$
    (c) Z_o
    (d) $Y^2 + R^2$
17. Inductive susceptance is defined in
    (a) imaginary ohms.
    (b) imaginary henrys.
    (c) imaginary farads.
    (d) imaginary siemens.
18. Capacitive susceptance values can be defined by
    (a) positive real numbers.
    (b) negative real numbers.
    (c) positive imaginary numbers.
    (d) negative imaginary numbers.
19. Which of the following is false?
    (a) $B_C = 1/X_C$.
    (b) Complex impedance can be depicted as a vector.
    (c) Characteristic impedance is complex.
    (d) $G = 1/R$.
20. In general, as the absolute value of the impedance in a circuit increases,
    (a) the flow of ac increases.
    (b) the flow of ac decreases.
    (c) the reactance decreases.
    (d) the resistance decreases.

16

CHAPTER

RLC and *GLC* Circuit Analysis

WHEN YOU SEE AN AC CIRCUIT THAT CONTAINS COILS AND/OR CAPACITORS, YOU SHOULD ENVISION a complex-number plane, either *RX* (resistance-reactance) or *GB* (conductance-admittance). The *RX* plane applies to series circuit analysis. The *GB* plane applies to parallel circuit analysis.

Complex Impedances in Series

When you see resistors, coils, and capacitors in series, each component has an impedance that can be represented as a vector in the *RX* plane. The vectors for resistors are constant, regardless of the frequency. But the vectors for coils and capacitors vary with frequency.

Pure Reactances

Pure inductive reactances (X_L) and capacitive reactances (X_C) simply add together when coils and capacitors are in series. Thus, $X = X_L + X_C$. In the *RX* plane, their vectors add, but because these vectors point in exactly opposite directions—inductive reactance upward and capacitive reactance downward (Fig. 16-1)—the resultant sum vector inevitably points either straight up or straight down, unless the reactances are equal and opposite, in which case they cancel and the result is the zero vector.

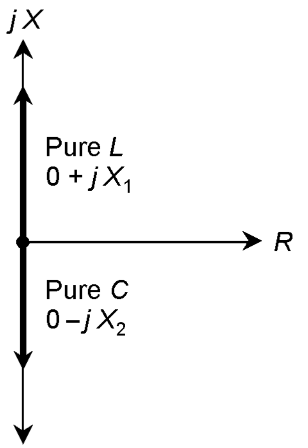
Problem 16-1

Suppose a coil and capacitor are connected in series, with $jX_L = j200$ and $jX_C = -j150$. What is the net reactance?

Just add the values: $jX = jX_L + jX_C = j200 + (-j150) = j(200 - 150) = j50$. This is a pure inductive reactance, because it is positive imaginary.

Problem 16-2

Suppose a coil and capacitor are connected in series, with $jX_L = j30$ and $jX_C = -j110$. What is the net reactance?



16-1 Pure inductance and pure capacitance are represented by reactance vectors that point straight up and down.

Again, add the values: $jX = j30 + (-j110) = j(30 - 110) = -j80$. This is a pure capacitive reactance, because it is negative imaginary.

Problem 16-3

Suppose a coil of inductance $L = 5.00$ μH and a capacitor of capacitance $C = 200$ pF are connected in series. Suppose the frequency is $f = 4.00$ MHz. What is the net reactance?

First, calculate the reactance of the inductor at 4.00 MHz. Proceed as follows:

$$\begin{aligned} jX_L &= j6.28fL \\ &= j(6.28 \times 4.00 \times 5.00) \\ &= j126 \end{aligned}$$

Next, calculate the reactance of the capacitor at 4.00 MHz. Proceed as follows:

$$\begin{aligned} jX_C &= -j[1/(6.28fC)] \\ &= -j[1/(6.28 \times 4.00 \times 0.000200)] \\ &= -j199 \end{aligned}$$

Finally, add the inductive and capacitive reactances to obtain the net reactance:

$$\begin{aligned} jX &= jX_L + jX_C \\ &= j126 + (-j199) \\ &= -j73 \end{aligned}$$

This is a pure capacitive reactance.

Problem 16-4

What is the net reactance of the aforementioned inductor and capacitor combination at the frequency $f = 10.0$ MHz?

First, calculate the reactance of the inductor at 10.0 MHz. Proceed as follows:

$$\begin{aligned} jX_L &= j6.28fL \\ &= j(6.28 \times 10.0 \times 5.00) \\ &= j314 \end{aligned}$$

Next, calculate the reactance of the capacitor at 10.00 MHz. Proceed as follows:

$$\begin{aligned} jX_C &= -j[1/(6.28fC)] \\ &= -j[1/(6.28 \times 10.0 \times 0.000200)] \\ &= -j79.6 \end{aligned}$$

Finally, add the inductive and capacitive reactances to obtain the net reactance:

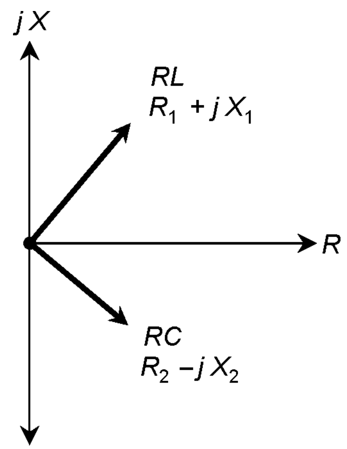
$$\begin{aligned} jX &= jX_L + jX_C \\ &= j314 + (-j79.6) \\ &= j234 \end{aligned}$$

This is a pure inductive reactance. For series-connected components, the condition in which the capacitive and inductive reactances cancel is known as *series resonance.* We'll deal with this in more detail in the next chapter.

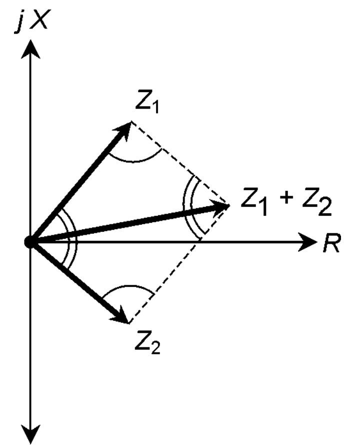
Adding Impedance Vectors

In the real world, there is resistance, as well as reactance, in an ac series circuit containing a coil and capacitor. This occurs because the coil wire has some resistance (it's never a perfect conductor). It can also be the case because a resistor is deliberately connected into the circuit.

Whenever the resistance in a series circuit is significant, the impedance vectors no longer point straight up and straight down. Instead, they run off toward the northeast (for the inductive part of the circuit) and southeast (for the capacitive part). This is illustrated in Fig. 16-2.



16-2 When resistance is present along with reactance, impedance vectors point at angles; they are neither vertical nor horizontal.



16-3 Parallelogram method of complex-impedance vector addition.

When two impedance vectors don't lie along a single line, you must use *vector addition* to be sure that you get the correct net impedance. In Fig. 16-3, the geometry of vector addition is shown. Construct a *parallelogram*, using the two vectors $Z_1 = R_1 + jX_1$ and $Z_2 = R_2 + jX_2$ as two adjacent sides of the figure. The diagonal of the parallelogram is the vector representing the net complex impedance. (Note that in a parallelogram, pairs of opposite angles have equal measures. These equalities are indicated by single and double arcs in Fig. 16-3.)

Formula for Complex Impedances in Series

Suppose you are given two complex impedances, $Z_1 = R_1 + jX_1$ and $Z_2 = R_2 + jX_2$. The net impedance, Z, of these in series is their vector sum, given by the following formula:

$$\begin{aligned} Z &= (R_1 + jX_1) + (R_2 + jX_2) \\ &= (R_1 + R_2) + j(X_1 + X_2) \end{aligned}$$

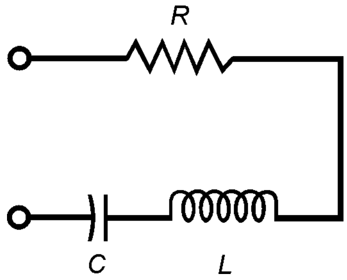
Calculating a vector sum using the formula is easier than doing it geometrically with a parallelogram. The arithmetic method is also more exact. The resistance and reactance components add separately. Just remember that if a reactance is capacitive, then it is negative imaginary in this formula.

Series *RLC* Circuits

When an inductance, capacitance, and resistance are connected in series (Fig. 16-4), the resistance R can be imagined as belonging entirely to the coil, when you use the preceding formulas. Then you have two vectors to add, when finding the impedance of the series *RLC* circuit containing three such components:

$$\begin{aligned} Z &= (R + jX_L) + (0 + jX_C) \\ &= R + j(X_L + X_C) \end{aligned}$$

Again, remember that X_C is never positive! So, although the formulas here have addition symbols in them, you're adding a negative number when you add in a capacitive reactance.



16-4 A series resistance-inductance-capacitance (*RLC*) circuit.

Problem 16-5

Suppose a resistor, a coil, and a capacitor are connected in series with $R = 50\ \Omega$, $X_L = 22\ \Omega$, and $X_C = -33\ \Omega$. What is the net impedance, Z?

Consider the resistor to be part of the coil, obtaining two complex vectors, $50 + j22$ and $0 - j33$. Adding these gives the resistance component of $50 + 0 = 50$, and the reactive component of $j22 - j33 = -j11$. Therefore, $Z = 50 - j11$.

Problem 16-6

Consider a resistor, a coil, and a capacitor that are connected in series with $R = 600\ \Omega$, $X_L = 444\ \Omega$, and $X_C = -444\ \Omega$. What is the net impedance, Z?

Again, imagine the resistor to be part of the inductor. Then the complex impedance vectors are $600 + j444$ and $0 - j444$. Adding these, the resistance component is $600 + 0 = 600$, and the reactive component is $j444 - j444 = j0$. Thus, $Z = 600 + j0$. This is a purely resistive impedance, and you can rightly call it $600\ \Omega$.

Problem 16-7

Suppose a resistor, a coil, and a capacitor are connected in series. The resistor has a value of $330\ \Omega$, the capacitance is 220 pF, and the inductance is 100 μH. The frequency is 7.15 MHz. What is the complex impedance of this series *RLC* circuit at this frequency?

First, calculate the inductive reactance. Remember that $X_L = 6.28fL$ and that megahertz and microhenrys go together in the formula. Multiply to obtain the following:

$$\begin{aligned} jX_L &= j(6.28 \times 7.15 \times 100) \\ &= j4490 \end{aligned}$$

Next, calculate the capacitive reactance using the formula $X_C = -1/(6.28fC)$. Convert 220 pF to microfarads to obtain $C = 0.000220$ μF. Then calculate:

$$\begin{aligned} jX_C &= -j[1/(6.28 \times 7.15 \times 0.000220)] \\ &= -j101 \end{aligned}$$

Now, lump the resistance and the inductive reactance together, so one of the impedance vectors is $330 + j4490$. The other is $0 - j101$. Adding these gives $Z = 330 + j4389$; this rounds off to $Z = 330 + j4390$.

Problem 16-8

Suppose a resistor, a coil, and a capacitor are connected in series. The resistance is 50.0 Ω, the inductance is 10.0 μH, and the capacitance is 1000 pF. The frequency is 1592 kHz. What is the complex impedance of this series *RLC* circuit at this frequency?

First, calculate $X_L = 6.28fL$. Convert the frequency to megahertz; 1592 kHz = 1.592 MHz. Then:

$$\begin{aligned} jX_L &= j(6.28 \times 1.592 \times 10.0) \\ &= j100 \end{aligned}$$

Then calculate $X_C = -1/(6.28fC)$. Let's convert picofarads to microfarads, and use megahertz for the frequency. Therefore:

$$\begin{aligned} jX_C &= -j[1/(6.28 \times 1.592 \times 0.001000)] \\ &= -j100 \end{aligned}$$

Let the resistance and inductive reactance go together as one vector, $50.0 + j100$. Let the capacitive reactance be represented as $0 - j100$. The sum is $Z = 50.0 + j100 - j100 = 50.0 + j0$. This is a pure resistance of 50.0 Ω. You can correctly say that the impedance is 50.0 Ω in this case.

Complex Admittances in Parallel

When you see resistors, coils, and capacitors in parallel, remember that each component, whether it is a resistor, an inductor, or a capacitor, has an admittance that can be represented as a vector in the *GB* plane. The vectors for pure conductances are constant, even as the frequency changes. But the vectors for the coils and capacitors vary with frequency.

Pure Susceptances

Pure inductive susceptances (B_L) and capacitive susceptances (B_C) add together when coils and capacitors are in parallel. Thus, $B = B_L + B_C$. Remember that B_L is never positive, and B_C is never negative. This is just the opposite situation from reactances.

In the *GB* plane, pure jB_L and jB_C vectors add. Because such vectors always point in exactly opposite directions—inductive susceptance down and capacitive susceptance up—the sum, jB, inevitably points either straight down or straight up (Fig. 16-5), unless the susceptances are equal and opposite, in which case they cancel and the result is the zero vector.

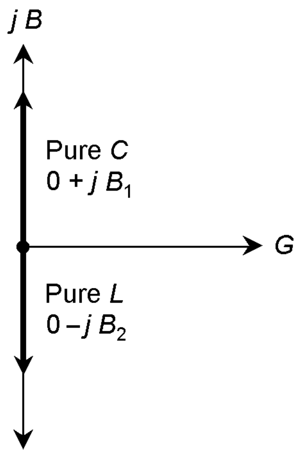
Problem 16-9

Suppose a coil and capacitor are connected in parallel, with $jB_L = -j0.05$ and $jB_C = j0.08$. What is the net susceptance?

Just add the values as follows: $jB = jB_L + jB_C = -j0.05 + j0.08 = j0.03$. This is a capacitive susceptance, because it is positive imaginary.

Problem 16-10

Suppose a coil and capacitor are connected in parallel, with $jB_L = -j0.60$ and $jB_C = j0.25$. What is the net susceptance?



16-5 Pure capacitance and pure inductance are represented by susceptance vectors that point straight up and down.

Again, add the values: $jB = -j0.60 + j0.25 = -j0.35$. This is an inductive susceptance, because it is negative imaginary.

Problem 16-11

Suppose a coil of $L = 6.00\ \mu\text{H}$ and a capacitor of $C = 150$ pF are connected in parallel. The frequency is $f = 4.00$ MHz. What is the net susceptance?

First calculate the susceptance of the inductor at 4.00 MHz, as follows:

$$\begin{aligned} jB_L &= -j[1/(6.28fL)] \\ &= -j[1/(6.28 \times 4.00 \times 6.00)] \\ &= -j0.00663 \end{aligned}$$

Next, calculate the susceptance of the capacitor (converting its value to microfarads) at 4.00 MHz, as follows:

$$\begin{aligned} jB_C &= j(6.28fC) \\ &= j(6.28 \times 4.00 \times 0.000150) \\ &= j0.00377 \end{aligned}$$

Finally, add the inductive and capacitive susceptances to obtain the net susceptance:

$$\begin{aligned} jB &= jB_L + jB_C \\ &= -j0.00663 + j0.00377 \\ &= -j0.00286 \end{aligned}$$

This is a pure inductive susceptance.

Problem 16-12

What is the net susceptance of the above parallel-connected inductor and capacitor at a frequency of $f = 5.31$ MHz?

First calculate the susceptance of the inductor at 5.31 MHz, as follows:

$$\begin{aligned} jB_L &= -j[1/(6.28fL)] \\ &= -j[1/(6.28 \times 5.31 \times 6.00)] \\ &= -j0.00500 \end{aligned}$$

Next calculate the susceptance of the capacitor (converting its value to microfarads) at 5.31 MHz, as follows:

$$\begin{aligned} jB_C &= \mathrm{j}(6.28fC) \\ &= j(6.28 \times 5.31 \times 0.000150) \\ &= j0.00500 \end{aligned}$$

Finally, add the inductive and capacitive susceptances to obtain the net susceptance:

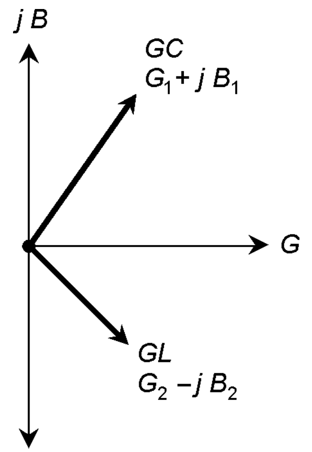
$$\begin{aligned} jB &= jB_L + jB_C \\ &= -j0.00500 + j0.00500 \\ &= j0 \end{aligned}$$

This means that the circuit has no susceptance at 5.31 MHz. The situation in which there is no susceptance in an *LC* circuit is known as *parallel resonance.* It is discussed in the next chapter.

Adding Admittance Vectors

In real life, there is a small amount of conductance, as well as susceptance, in an ac parallel circuit containing a coil and capacitor. This occurs when the capacitor lets a little bit of current leak through. More often, though, it is the case because a *load* is connected in parallel with the coil and capacitor. This load can be an antenna, the input to an amplifier circuit, a test instrument, a transducer, or some other device.

When the conductance in a parallel circuit containing inductance and capacitance is significant, the admittance vectors do not point straight up and down. Instead, they run off toward the northeast (for the capacitive part of the circuit) and southeast (for the inductive part). This is illustrated in Fig. 16-6.



16-6 When conductance is present along with susceptance, admittance vectors point at angles; they are neither vertical nor horizontal.

You've seen how vectors add in the *RX* plane. In the *GB* plane, the principle is the same. The net admittance vector is the sum of the component admittance vectors.

Formula for Complex Admittances in Parallel

Given two admittances, $Y_1 = G_1 + jB_1$ and $Y_2 = G_2 + jB_2$, the net admittance Y of these in parallel is their vector sum, as follows:

$$\begin{aligned} Y &= (G_1 + jB_1) + (G_2 + jB_2) \\ &= (G_1 + G_2) + j(B_1 + B_2) \end{aligned}$$

The conductance and susceptance components add separately. Just remember that if a susceptance is inductive, then it is negative imaginary in this formula.

Parallel *GLC* Circuits

When a coil, capacitor, and resistor are connected in parallel (Fig. 16-7), the resistance should be thought of as a *conductance*, whose value in siemens (symbolized S) is equal to the reciprocal of the value in ohms. Think of the conductance as all belonging to the inductor. Then you have two vectors to add, when finding the admittance of a parallel *GLC* (conductance-inductance-capacitance) circuit:

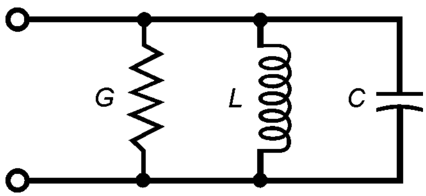
$$\begin{aligned} Y &= (G + jB_L) + (0 + jB_C) \\ &= G + j(B_L + B_C) \end{aligned}$$

Again, remember that B_L is never positive! So, although the formulas here have addition symbols in them, you're adding a negative number when you add in an inductive susceptance.

Problem 16-13

Suppose a resistor, a coil, and a capacitor are connected in parallel. Suppose the resistor has a conductance $G = 0.10$ S, and the susceptances are $jB_L = -j0.010$ and $jB_C = j0.020$. What is the complex admittance of this combination?

Consider the resistor to be part of the coil. Then there are two complex admittances in parallel: $0.10 - j0.010$ and $0.00 + j0.020$. Adding these gives a conductance component of $0.10 + 0.00 = 0.10$ and a susceptance component of $-j0.010 + j0.020 = j0.010$. Therefore, the complex admittance is $0.10 + j0.010$.



16-7 A parallel conductance-inductance-capacitance (*GLC*) circuit.

Problem 16-14

Suppose a resistor, a coil, and a capacitor are connected in parallel. Suppose the resistor has a conductance $G = 0.0010$ S, and the susceptances are $jB_L = -j0.0022$ and $jB_C = j0.0022$. What is the complex admittance of this combination?

Again, consider the resistor to be part of the coil. Then the complex admittances are $0.0010 - j0.0022$ and $0.0000 + j0.0022$. Adding these, the conductance component is $0.0010 + 0.0000 = 0.0010$, and the susceptance component is $-j0.0022 + j0.0022 = j0$. Thus, the admittance is $0.0010 + j0$. This is a purely conductive admittance.

Problem 16-15

Suppose a resistor, a coil, and a capacitor are connected in parallel. The resistor has a value of 100 Ω, the capacitance is 200 pF, and the inductance is 100 μH. The frequency is 1.00 MHz. What is the net complex admittance?

First, you need to calculate the inductive susceptance. Recall the formula, and plug in the numbers as follows:

$$\begin{aligned} jB_L &= -j[1/(6.28fL)] \\ &= -j[1/(6.28 \times 1.00 \times 100)] \\ &= -j0.00159 \end{aligned}$$

Megahertz and microhenrys go together in the formula. Next, you must calculate the capacitive susceptance. Convert 200 pF to microfarads to go with megahertz in the formula; thus $C = 0.000200$ μF. Then:

$$\begin{aligned} jB_C &= j(6.28fC) \\ &= j(6.28 \times 1.00 \times 0.000200) \\ &= j0.00126 \end{aligned}$$

Finally, consider the conductance, which is $1/100 = 0.0100$ S, and the inductive susceptance as existing together in a single component. That means that one of the parallel-connected admittances is $0.0100 - j0.00159$. The other is $0.0000 + j0.00126$. Adding these gives $0.0100 - j0.00033$.

Problem 16-16

Suppose a resistor, a coil, and a capacitor are in parallel. The resistance is 10.0 Ω, the inductance is 10.0 μH, and the capacitance is 1000 pF. The frequency is 1592 kHz. What is the complex admittance of this circuit at this frequency?

First, calculate the inductive susceptance. Convert the frequency to megahertz; 1592 kHz = 1.592 MHz. Plug in the numbers as follows:

$$\begin{aligned} jB_L &= -j[1/(6.28fL)] \\ &= -j[1/(6.28 \times 1.592 \times 10.0)] \\ &= -j0.0100 \end{aligned}$$

Next, calculate the capacitive susceptance. Convert 1000 pF to microfarads to go with megahertz in the formula; thus $C = 0.001000$ μF. Then:

$$\begin{aligned} jB_C &= j(6.28fC) \\ &= j(6.28 \times 1.592 \times 0.001000) \\ &= j0.0100 \end{aligned}$$

Finally, consider the conductance, which is 1/10.0 = 0.100 S, and the inductive susceptance as existing together in a single component. That means that one of the parallel-connected admittances is $0.100 - j0.0100$. The other is $0.0000 + j0.0100$. Adding these gives $0.100 + j0$.

Converting Complex Admittance to Complex Impedance

The *GB* plane is, as you have seen, similar in appearance to the *RX* plane, although mathematically they are different. Once you've found a complex admittance for a parallel *RLC* circuit, you will usually want to transform this back to a complex impedance.

The transformation from a complex admittance $G + jB$ to a complex impedance $R + jX$ can be carried out using the following two formulas, one for R and the other for X:

$$\begin{aligned} R &= G/(G^2 + B^2) \\ X &= -B/(G^2 + B^2) \end{aligned}$$

If you know the complex admittance, first find the resistance and reactance components individually using the preceding formulas. Then assemble the two components into the complex impedance, $R + jX$.

Problem 16-17

Suppose the complex admittance of a certain parallel circuit is $0.010 - j0.0050$. What is the complex impedance of this same circuit, assuming the frequency does not change?

In this case, $G = 0.010$ S and $B = -0.0050$ S. First find $G^2 + B^2$, as follows:

$$\begin{aligned} G^2 + B^2 &= 0.010^2 + (-0.0050)^2 \\ &= 0.000100 + 0.000025 \\ &= 0.000125 \end{aligned}$$

Now it is easy to calculate R and X, like this:

$$\begin{aligned} R &= G/0.000125 \\ &= 0.010/0.000125 \\ &= 80\ \Omega \end{aligned}$$

$$\begin{aligned} X &= -B/0.000125 \\ &= 0.0050/0.000125 \\ &= 40\ \Omega \end{aligned}$$

The complex impedance is therefore $80 + j40$.

Putting It All Together

When you're confronted with a parallel circuit containing resistance, inductance, and capacitance, and you want to determine the complex impedance of the combination, do these things:

1. Find the conductance $G = 1/R$ for the resistor. (It will be positive or zero.)
2. Find the susceptance B_L of the inductor using the appropriate formula. (It will be negative or zero.)
3. Find the susceptance B_C of the capacitor using the appropriate formula. (It will be positive or zero.)
4. Find the net susceptance $B = B_L + B_C$. (It might be positive, negative, or zero.)
5. Compute R and X in terms of G and B using the appropriate formulas.
6. Assemble the complex impedance $R + jX$.

Problem 16-18

Suppose a resistor of 10.0 Ω, a capacitor of 820 pF, and a coil of 10.0 μH are in parallel. The frequency is 1.00 MHz. What is the complex impedance?

Proceed according to the above steps, as follows:

1. Calculate $G = 1/R = 1/10.0 = 0.100$.
2. Calculate $B_L = -1/(6.28fL) = -1/(6.28 \times 1.00 \times 10.0) = -0.0159$.
3. Calculate $B_C = 6.28fC = 6.28 \times 1.00 \times 0.000820 = 0.00515$. (Remember to first convert the capacitance to microfarads, to go with megahertz.)
4. Calculate $B = B_L + B_C = -0.0159 + 0.00515 = -0.0108$.
5. Define $G^2 + B^2 = 0.100^2 + (-0.0108)^2 = 0.010117$. Then $R = G/0.010117 = 0.100/0.010117 = 9.88\ \Omega$, and $X = -B/0.010117 = 0.0108/0.010117 = 1.07\ \Omega$.
6. The complex impedance is $R + jX = 9.88 + j1.07$.

Problem 16-19

Suppose a resistor of 47.0 Ω, a capacitor of 500 pF, and a coil of 10.0 μH are in parallel. What is their complex impedance at a frequency of 2.252 MHz?

Proceed as before:

1. Calculate $G = 1/R = 1/47.0 = 0.021277$.
2. Calculate $B_L = -1/(6.28fL) = -1/(6.28 \times 2.252 \times 10.0) = -0.00707$.
3. Calculate $B_C = 6.28fC = 6.28 \times 2.252 \times 0.000500 = 0.00707$. (Remember to first convert the capacitance to microfarads, to go with megahertz.)
4. Calculate $B = B_L + B_C = -0.00707 + 0.00707 = 0.00000$.
5. Define $G^2 + B^2 = 0.021277^2 + 0.00000^2 = 0.00045271$. Then $R = G/0.00045271 = 0.021277/0.00045271 = 46.999\ \Omega$, and $X = -B/0.00045271 = 0.00000/0.00045271 = 0.00000$.
6. The complex impedance is $R + jX = 46.9999 + j0.00000$. When we round it off to three significant figures, we get $47.0 + j0.00$. This a pure resistance equal to the value of the resistor in the circuit.

Reducing Complicated *RLC* Circuits

Sometimes you'll see circuits in which there are several resistors, capacitors, and/or coils in series and parallel combinations. Such a circuit can be reduced to an equivalent series or parallel *RLC* circuit that contains one resistance, one capacitance, and one inductance.

Series Combinations

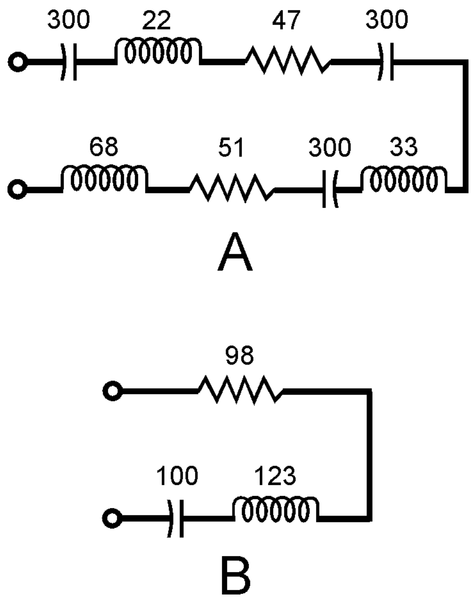
Resistances in series simply add. Inductances in series also add. Capacitances in series combine in a somewhat more complicated way. If you don't remember the formula, here it is:

$$1/C = 1/C_1 + 1/C_2 + \cdots + 1/C_n$$

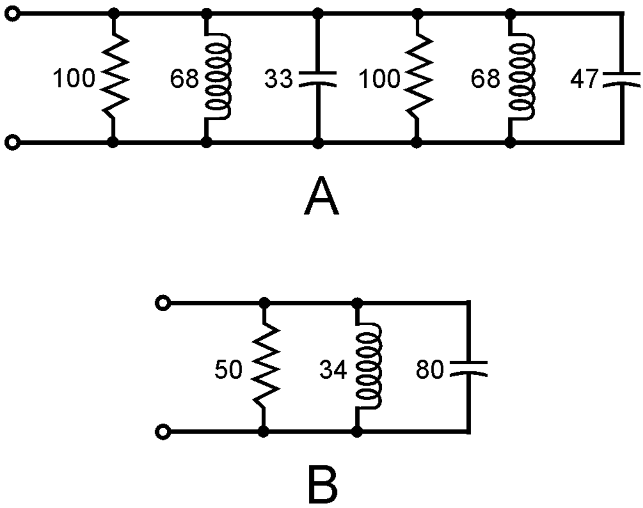
where C_1, C_2, . . . , and C_n are the individual capacitances, and C is the total capacitance. Once you've found $1/C$, take its reciprocal to obtain C. Figure 16-8A shows an example of a complicated series *RLC* circuit. The equivalent circuit, with one resistance, one capacitance, and one inductance, is shown in Fig. 16-8B.

Parallel Combinations

In parallel, resistances and inductances combine the way capacitances do in series. Capacitances simply add up. An example of a complicated parallel *RLC* circuit is shown in Fig. 16-9A. The equivalent circuit, with one resistance, one capacitance, and one inductance, is shown in Fig. 16-9B.



16-8 At A, a complicated series circuit containing multiple resistances and reactances. At B, the same circuit simplified. Resistances are in ohms; inductances are in microhenrys (μH); capacitances are in picofarads (pF).

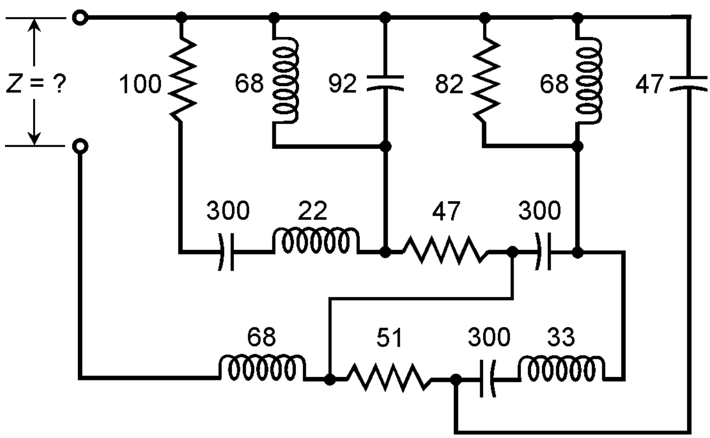


16-9 At A, a complicated parallel circuit containing multiple resistances and reactances. At B, the same circuit simplified. Resistances are in ohms; inductances are in microhenrys (μH); capacitances are in picofarads (pF).

Nightmare Scenarios

Imagine an *RLC* circuit like the one shown in Fig. 16-10. How would you find the complex impedance of this circuit at some particular frequency, such as 8.54 MHz? Don't waste much time worrying about circuits like this. You'll rarely encounter them. But rest assured that, given a frequency, a complex impedance does exist, no matter how complicated an *RLC* circuit happens to be.

An engineer could use a computer to find the theoretical complex impedance of a circuit such as the one in Fig. 16-10 at a specific frequency, or as a function of the frequency. The experimental approach would be to build the circuit, connect a signal generator to it, and then measure *R* and *X* at various frequencies with a device called an *impedance bridge.*



16-10 A series-parallel nightmare circuit containing multiple resistances and reactances. Resistances are in ohms; inductances are in microhenrys (μH); capacitances are in picofarads (pF).

Ohm's Law for AC Circuits

Ohm's Law for a dc circuit is a simple relationship among three variables: the current I (in amperes), the voltage E (in volts), and the resistance R (in ohms). Here are the formulas, in case you don't recall them:

$$E = IR$$
$$I = E/R$$
$$R = E/I$$

In ac circuits containing no reactance, these same formulas apply, as long as you work with root-mean-square (rms) voltages and currents. If you need a refresher concerning the meaning of rms, refer to Chapter 9.

Purely Resistive Impedances

When the impedance Z in an ac circuit contains no reactance, so that all of the current and voltage exist through and across a pure resistance R, Ohm's Law for an ac circuit is expressed as follows:

$$E = IZ$$
$$I = E/Z$$
$$Z = E/I$$

where $Z = R$, and the values I and E are rms current and voltage.

Complex Impedances

When you want to determine the relationship among current, voltage, and resistance in an ac circuit that contains resistance and reactance, things get interesting. Recall the formula for the square of the absolute-value impedance in a series *RLC* circuit:

$$Z^2 = R^2 + X^2$$

This means that Z is equal to the square root of the quantity $R^2 + X^2$, as follows:

$$Z = (R^2 + X^2)^{1/2}$$

This is the length of the vector $R + jX$ in the complex impedance plane. You learned this in Chap. 15. This formula applies only for series *RLC* circuits.

The square of the absolute-value impedance for a parallel *RLC* circuit, in which the resistance is R and the reactance is X, is defined this way:

$$Z^2 = R^2X^2/(R^2 + X^2)$$

This means that the absolute-value impedance, Z, must be calculated using the rather arcane formula:

$$Z = [R^2X^2/(R^2 + X^2)]^{1/2}$$

The $\frac{1}{2}$ power of a quantity represents the positive square root of that quantity.

Problem 16-20

Suppose a series *RX* circuit (shown by the generic block diagram of Fig. 16-11) has a resistance of $R = 50.0\ \Omega$ and a capacitive reactance of $X = -50.0\ \Omega$. Suppose 100-V rms ac is applied to this circuit. What is the current?

First, calculate $Z^2 = R^2 + X^2 = 50.0^2 + (-50.0)^2 = 2500 + 2500 = 5000$. Then Z is the square root of 5000, or 70.7. Therefore, $I = E/Z = 100/70.7 = 1.41$ A rms.

Problem 16-21

What are the rms ac voltages across the resistance and the reactance, respectively, in the circuit described in Problem 16-20?

The Ohm's Law formulas for dc will work here. Because the current is $I = 1.41$ A rms, the voltage drop across the resistance is equal to $E_R = IR = 1.41 \times 50.0 = 70.5$ V rms. The voltage drop across the reactance is the product of the current and the reactance: $E_X = IX = 1.41 \times (-50.0) = -70.5$ V rms. This is an rms ac voltage of equal magnitude to that across the resistance. But the phase is different.

Note that voltages across the resistance and the reactance—a capacitive reactance in this case, because it's negative—don't add up to 100 V rms, which is placed across the whole circuit. This is because, in an *RX* ac circuit, there is always a difference in phase between the voltage across the resistance and the voltage across the reactance. The voltages across the components always add up to the applied voltage *vectorially*, but not always *arithmetically.*

Problem 16-22

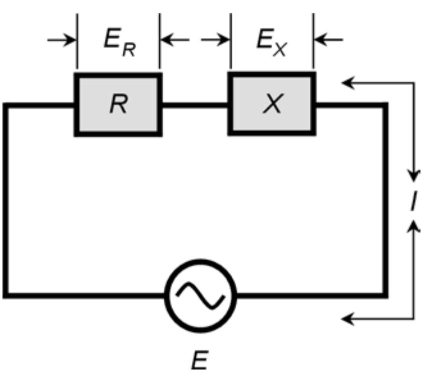
Suppose a series *RX* circuit (Fig. 16-11) has $R = 10.0\ \Omega$. and $X = 40.0\ \Omega$. The applied voltage is 100-V rms ac. What is the current?

Calculate $Z^2 = R^2 + X^2 = 100 + 1600 = 1700$. This means that Z is the square root of 1700, or 41.2. Therefore, $I = E/Z = 100/41.2 = 2.43$ A rms.

Problem 16-23

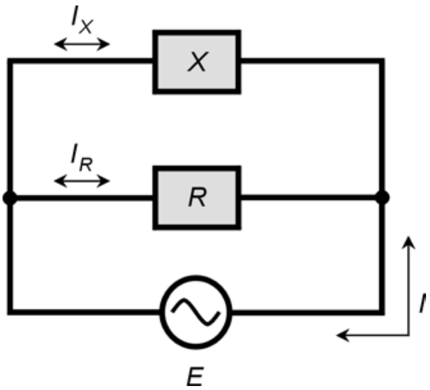
What are the rms ac voltages across the resistance and the reactance, respectively, in the circuit described in Problem 16-22?

Knowing the current, calculate $E_R = IR = 2.43 \times 10.0 = 24.3$ V rms. Also, $E_X = IX = 2.43 \times 40.0 = 97.2$ V rms. If you add $E_R + E_X$ arithmetically, you get $24.3 + 97.2 = 121.5$ V as the total



16-11 A series circuit containing resistance and reactance. Illustration for Problems 16-20 through 16-23.

16-12 A parallel circuit containing resistance and reactance. Illustration for Problems 16-24 and 16-25.



across R and X. Again, this differs from the applied voltage! The simple dc rule does not work here, for the same reason it didn't work in the scenario of Problem 16-21.

Problem 16-24

Suppose a parallel RX circuit (shown by the generic block diagram of Fig. 16-12) has $R = 30.0\ \Omega$ and $X = -20.0\ \Omega$. The ac supply voltage is 50.0 V rms. What is the total current drawn from the ac supply?

First, find the square of the absolute-value impedance, remembering the formula for parallel circuits: $Z^2 = R^2X^2/(R^2 + X^2) = 360{,}000/1300 = 277$. The absolute-value impedance Z is the square root of 277, or 16.6. The total current is therefore $I = E/Z = 50/16.6 = 3.01$ A rms.

Problem 16-25

What are the rms currents through the resistance and the reactance, respectively, in the circuit described in Problem 16-24?

The Ohm's Law formulas for dc will work here. For the resistance, $I_R = E/R = 50.0/30.0 = 1.67$ A rms. For the reactance, $I_X = E/X = 50.0/(-20.0) = -2.5$ A rms. Note that these currents don't add up to 3.01 A, the total current. The reason for this is the same as the reason ac voltages don't add arithmetically in ac circuits that contain reactance. The constituent currents, I_R and I_X, differ in phase. Vectorially, they add up to 3.01 A rms, but arithmetically, they don't.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 correct. Answers are in the back of the book.

1. Suppose a coil and capacitor are connected in series. The inductive reactance is 250 Ω, and the capacitive reactance is −300 Ω. What is the complex impedance?
   (a) $0 + j550$
   (b) $0 - j50$
   (c) $250 - j300$
   (d) $-300 + j250$

2. Suppose a coil of 25.0 μH and capacitor of 100 pF are connected in series. The frequency is 5.00 MHz. What is the complex impedance?
   (a) $0 + j467$
   (b) $25 + j100$
   (c) $0 - j467$
   (d) $25 - j100$
3. When $R = 0$ in a series *RLC* circuit, but the net reactance is not zero, the impedance vector
   (a) always points straight up.
   (b) always points straight down.
   (c) always points straight toward the right.
   (d) None of the above is correct.
4. Suppose a resistor of 150 Ω, a coil with a reactance of 100 Ω, and a capacitor with a reactance of −200 Ω are connected in series. What is the complex impedance?
   (a) $150 + j100$
   (b) $150 - j200$
   (c) $100 - j200$
   (d) $150 - j100$
5. Suppose a resistor of 330 Ω, a coil of 1.00 μH, and a capacitor of 200 pF are in series. What is the complex impedance at 10.0 MHz?
   (a) $330 - j199$
   (b) $300 + j201$
   (c) $300 + j142$
   (d) $330 - j16.8$
6. Suppose a coil has an inductance of 3.00 μH and a resistance of 10.0 Ω in its winding. A capacitor of 100 pF is in series with this coil. What is the complex impedance at 10.0 MHz?
   (a) $10 + j3.00$
   (b) $10 + j29.2$
   (c) $10 - j97$
   (d) $10 + j348$
7. Suppose a coil has a reactance of 4.00 Ω. What is the complex admittance, assuming there is nothing else is in the circuit?
   (a) $0 + j0.25$
   (b) $0 + j4.00$
   (c) $0 - j0.25$
   (d) $0 - j4.00$
8. What will happen to the susceptance of a capacitor if the frequency is doubled and all other factors remain constant?
   (a) It will decrease to half its former value.
   (b) It will not change.

(c) It will double.

(d) It will quadruple.

9. Suppose a coil and capacitor are in parallel, with $jB_L = -j0.05$ and $jB_C = j0.03$. What is the complex admittance, assuming that nothing is in series or parallel with these components?

(a) $0 - j0.02$

(b) $0 - j0.07$

(c) $0 + j0.02$

(d) $-0.05 + j0.03$

10. Imagine a coil, a resistor, and a capacitor connected in parallel. The resistance is 1.0 Ω, the capacitive susceptance is 1.0 S, and the inductive susceptance is −1.0 S. Then, suddenly, the frequency is cut to half its former value. What is the complex admittance at the new frequency?

(a) $1.0 + j0.0$

(b) $1.0 + j1.5$

(c) $1.0 - j1.5$

(d) $1.0 - j2.0$

11. Suppose a coil of 3.50 μH and a capacitor of 47.0 pF are in parallel. The frequency is 9.55 MHz. There is nothing else in series or parallel with these components. What is the complex admittance?

(a) $0 + j0.00282$

(b) $0 - j0.00194$

(c) $0 + j0.00194$

(d) $0 - j0.00758$

12. A vector pointing southeast in the *GB* plane would indicate

(a) pure conductance with zero susceptance.

(b) conductance and inductive susceptance.

(c) conductance and capacitive susceptance.

(d) pure susceptance with zero conductance.

13. Suppose a resistor with conductance 0.0044 S, a capacitor with susceptance 0.035 S, and a coil with susceptance −0.011 S are all connected in parallel. What is the complex admittance?

(a) $0.0044 + j0.024$

(b) $0.035 - j0.011$

(c) $-0.011 + j0.035$

(d) $0.0044 + j0.046$

14. Suppose a resistor of 100 Ω, a coil of 4.50 μH, and a capacitor of 220 pF are in parallel. What is the complex admittance at a frequency of 6.50 MHz?

(a) $100 + j0.00354$

(b) $0.010 + j0.00354$

(c) $100 - j0.0144$

(d) $0.010 + j0.0144$

15. Suppose the complex admittance of a circuit is $0.02 + j0.20$. What is the complex impedance, assuming the frequency does not change?
    (a) $50 + j5.0$
    (b) $0.495 - j4.95$
    (c) $50 - j5.0$
    (d) $0.495 + j4.95$

16. Suppose a resistor of 51.0 Ω, an inductor of 22.0 μH, and a capacitor of 150 pF are in parallel. The frequency is 1.00 MHz. What is the complex impedance?
    (a) $51.0 - j14.9$
    (b) $51.0 + j14.9$
    (c) $46.2 - j14.9$
    (d) $46.2 + j14.9$

17. Suppose a series circuit has 99.0 Ω of resistance and 88.0 Ω of inductive reactance. An ac rms voltage of 117 V is applied to this series network. What is the current?
    (a) 1.18 A
    (b) 1.13 A
    (c) 0.886 A
    (d) 0.846 A

18. What is the voltage across the reactance in the preceding example?
    (a) 78.0 V
    (b) 55.1 V
    (c) 99.4 V
    (d) 74.4 V

19. Suppose a parallel circuit has 10 Ω of resistance and 15 Ω of reactance. An ac rms voltage of 20 V is applied across it. What is the total current?
    (a) 2.00 A
    (b) 2.40 A
    (c) 1.33 A
    (d) 0.800 A

20. What is the current through the resistance in the preceding example?
    (a) 2.00 A
    (b) 2.40 A
    (c) 1.33 A
    (d) 0.800 A

17

CHAPTER

Power and Resonance in Alternating-Current Circuits

ONE OF THE BIGGEST CHALLENGES IN ELECTRICITY AND ELECTRONICS IS OPTIMIZING THE EFFICIENCY with which power is transferred from one place to another, or converted from one form to another. Also important, especially for the radio-frequency (RF) engineer, is the phenomenon of *resonance.* Power and resonance are closely related.

Forms of Power

What is power, exactly? Here is an all-encompassing definition: *Power is the rate at which energy is expended, radiated, or dissipated.* This definition can be applied to mechanical motion, chemical effects, dc and ac electricity, sound waves, radio waves, sound, heat, infrared (IR), visible light, ultraviolet (UV), X rays, gamma rays, and high-speed subatomic particles. In all cases, the energy is converted from one form into another form at a certain rate.

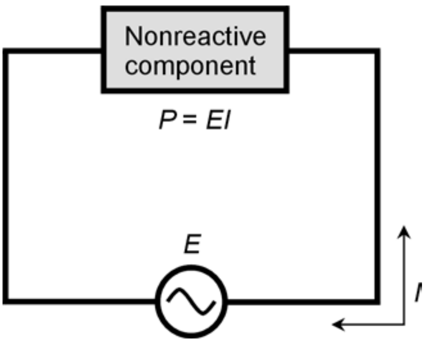
Units of Power

The standard unit of power is the *watt*, abbreviated W. A watt is equivalent to a *joule per second* (J/s). Sometimes power is given as *kilowatts* (kW or thousands of watts), *megawatts* (MW or millions of watts), or *gigawatts* (GW or billions of watts). It is also sometimes expressed as *milliwatts* (mW or thousandths of watts), *microwatts* (μW or millionths of watts), or *nanowatts* (nW or billionths of watts).

Volt-Amperes

In dc circuits, and also in ac circuits having no reactance, power can be defined this way: *Power is the product of the voltage across a circuit or component and the current through that same circuit or component.* Mathematically this is written $P = EI$. If E is in volts and I is in amperes, then P is in *volt-amperes* (VA). This translates into watts when there is no reactance in the circuit (Fig. 17-1). The root-mean-square (rms) values for voltage and current are always used to derive the effective, or average, power.

Volt-amperes, also called *VA power* or *apparent power*, can take various forms. A resistor converts electrical energy into heat energy, at a rate that depends on the value of the resistance and the current through it. A light bulb converts electricity into light and heat. A radio antenna converts high-



17-1 When there is no reactance in an ac component, the power *P* is the product of the voltage *E* across the component and the current *I* through the component.

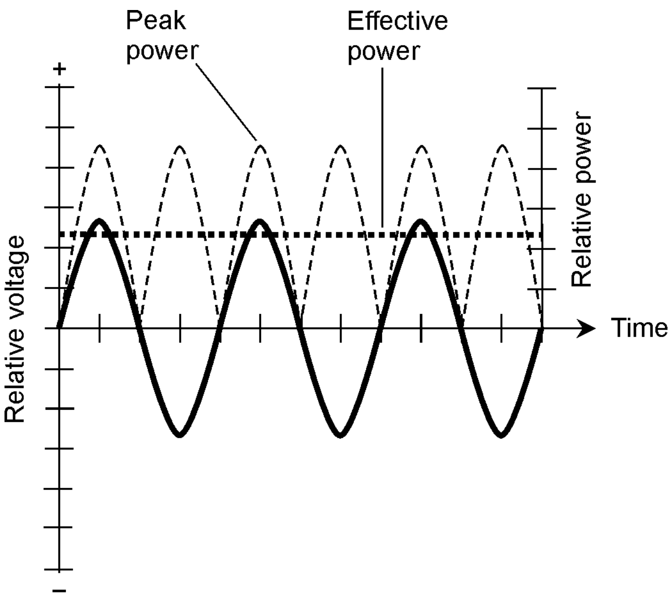
frequency ac into radio waves. A speaker converts low-frequency ac into sound waves. The power in these forms is a measure of the intensity of the heat, light, radio waves, or sound waves.

Instantaneous Power

Usually, but not always, engineers think of power based on the rms, or effective, ac value. But for VA power, peak values are sometimes used instead. If the ac is a sine wave, the peak current is 1.414 times the rms current, and the peak voltage is 1.414 times the rms voltage. If the current and the voltage are exactly in phase, the product of their peak values is twice the product of their rms values.

There are instants in time when the VA power in a reactance-free, sine-wave ac circuit is twice the effective power. There are other instants in time when the VA power is zero; at still other moments, the VA power is somewhere between zero and twice the effective power level (Fig. 17-2). This constantly changing power is called *instantaneous power.*

In some situations, such as with a voice-modulated radio signal or a fast-scan television signal, the instantaneous power varies in an extremely complicated fashion. Have you ever seen the *modulation envelope* of such a signal displayed on an oscilloscope?



17-2 Peak versus effective power for a sine wave. The left-hand vertical scale shows relative voltage. The right-hand vertical scale shows relative power. The solid curve represents the voltage as a function of time. The light and heavy dashed waves show peak and effective power, respectively, as functions of time.

Imaginary Power

If an ac circuit contains reactance, things get interesting. In a pure resistance, the rate of energy expenditure per unit time (or *true power*) is the same as the VA power (also known as *apparent power*). But when inductance and/or capacitance exists in an ac circuit, the VA power is greater than the power actually manifested as heat, light, radio waves, or whatever. The apparent power is then greater than the true power! The extra power is called *imaginary power*, because it exists in the reactance, and reactance can be, as you have learned, rendered in mathematically imaginary numerical form. Imaginary power is also known as *reactive power.*

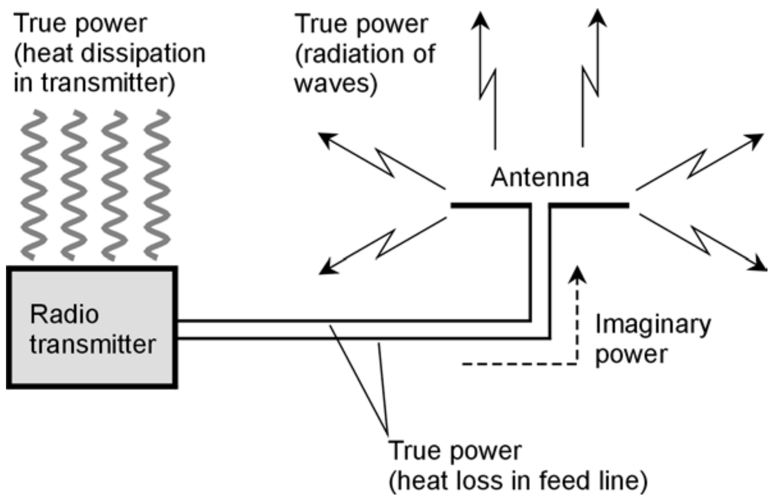
Inductors and capacitors store energy and then release it a fraction of a cycle later. This phenomenon, like true power, is expressible as the rate at which energy is changed from one form to another. But rather than existing as a usable form of power, such as heat, light, radio waves, sound waves, or mechanical motion, imaginary power is stored up as a magnetic or electric field, and then released back into the circuit or system. This storage and release of power takes place over and over with each repeating ac cycle.

True Power Does Not Travel

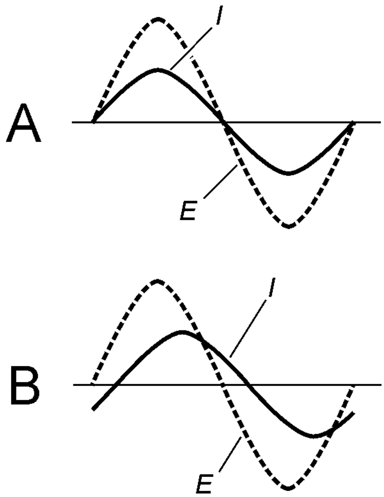
A common and usually harmless misconception about true power is the notion that it can travel. For example, if you connect a radio transmitter to a cable that runs outdoors to an antenna, you might say you're "feeding power" through the cable to the antenna. Everybody says this, even engineers and technicians. But true power always involves a change in form, such as from electrical current and voltage into radio waves. It doesn't go from place to place. It simply happens in a specific place. It's the imaginary power that moves in situations like this, especially in transmission lines between power stations and power users, or between radio transmitters and radio antennas.

In a real-life radio antenna system, some true power is dissipated as heat in the transmitter amplifiers and in the feed line (Fig. 17-3). The useful dissipation of true power occurs when the imaginary power, in the form of electric and magnetic fields, gets to the antenna, where it is changed into *electromagnetic waves.*

You will often hear expressions such as "forward power" and "reflected power," or "power is fed from this amplifier to these speakers." It is all right to talk like this, but it can sometimes lead to



17-3 True power and imaginary power in a radio transmitter and antenna system.



17-4 At A, current (I) and voltage (E) are in phase in a nonreactive ac circuit. At B, I and E are not in phase when reactance is present.

wrong conclusions, especially concerning impedance and standing waves. Then, you need to be keenly aware of the distinction among true, imaginary, and apparent power.

Reactance Does Not Consume Power

A pure inductance or a pure capacitance cannot dissipate any power. The only thing that such a component can do is store energy and then give it back to the circuit a fraction of a cycle later. In real life, the dielectrics or wires in coils and capacitors dissipate some power as heat, but ideal components would not do this.

A capacitor, as you have learned, stores energy as an electric field. An inductor stores energy as a magnetic field.

A component that contains reactance causes ac to shift in phase, so that the current is no longer exactly in step with the voltage. In a circuit with inductive reactance, the current lags the voltage by up to 90°, or one-quarter cycle. In a circuit with capacitive reactance, the current leads the voltage by up to 90°.

In a resistance-reactance circuit, true power is dissipated only in the resistive components. The reactive components exaggerate the VA power compared with the true power. Why, you ask, does reactance cause this discrepancy? In a circuit that is purely resistive, the voltage and current march right along in step with each other, and therefore, they combine in the most efficient possible way (Fig. 17-4A). But in a circuit containing reactance, the voltage and current are out of step with each other (Fig. 17-4B) because of their phase difference. Therefore, the actual energy expenditure, or true power, is not as great as the product of the voltage and the current.

True Power, VA Power, and Reactive Power

In an ac circuit or system containing nonzero resistance and nonzero reactance, the relationships among true power P_T, apparent (VA) power P_{VA}, and imaginary (reactive) power P_X are as follows:

$$P_{VA}^2 = P_T^2 + P_X^2$$
$$P_T < P_{VA}$$
$$P_X < P_{VA}$$

If there is no reactance in the circuit or system, then $P_{VA} = P_T$, and $P_X = 0$. Engineers strive to minimize, and if possible eliminate, the reactance in power-transmission systems.

Power Factor

In an ac circuit, the ratio of the true power to the VA power, P_T/P_{VA}, is called the *power factor.* If there is no reactance, the ideal case, then $P_T = P_{VA}$, and the power factor (PF) is equal to 1. If the circuit contains all reactance and no resistance of any significance (that is, zero or infinite resistance), then $P_T = 0$, and therefore $PF = 0$.

When a *load*, or a circuit in which you want power to be dissipated, contains resistance and reactance, then PF is between 0 and 1. That is, $0 < PF < 1$. The power factor can also be expressed as a percentage between 0 and 100, written $PF_{\%}$. Mathematically, we have these formulas for the power factor:

$$PF = P_T/P_{VA}$$
$$PF_{\%} = 100P_T/P_{VA}$$

When a load has some resistance and some reactance, then some of the power is dissipated as true power, and some is rejected by the load as imaginary power. In a sense, this imaginary power is sent back to the power source.

There are two ways to determine the power factor in an ac circuit that contains reactance and resistance. One method is to find the cosine of the phase angle. The other method involves the ratio of the resistance to the absolute-value impedance.

Cosine of Phase Angle

Recall that in a circuit having reactance and resistance, the current and the voltage are not in phase. The phase angle (ϕ) is the extent, expressed in degrees, to which the current and the voltage differ in phase. If there is no reactance, then $\phi = 0°$. If there is a pure reactance, then either $\phi = +90°$ (if the reactance is inductive) or else $\phi = -90°$ (if the reactance is capacitive). The power factor is equal to the cosine of the phase angle:

$$PF = \cos \phi$$

Problem 17-1

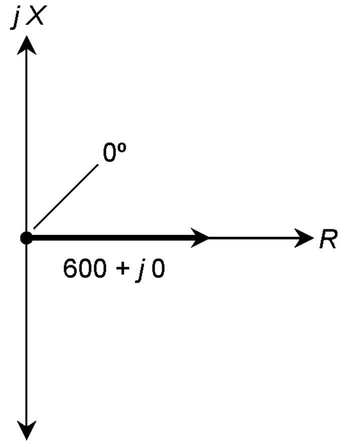
Suppose a circuit contains no reactance, but a pure resistance of 600 Ω. What is the power factor?

Without doing any calculations, it is evident that $PF = 1$, because $P_{VA} = P_T$ in a pure resistance. That means $P_T/P_{VA} = 1$. But you can also look at this by noting that the phase angle is 0°, because the current is in phase with the voltage. Using your calculator, you can see that $\cos 0° = 1$. Therefore, $PF = 1 = 100\%$. The vector for this case is shown in Fig. 17-5.

Problem 17-2

Suppose a circuit contains a pure capacitive reactance of −40 Ω, but no resistance. What is the power factor?

Here, the phase angle is −90° (Fig. 17-6). A calculator will tell you that $\cos -90° = 0$. Therefore, $PF = 0$, and $P_T/P_{VA} = 0 = 0\%$. None of the power is true; all of it is reactive.



17-5 Vector diagram showing the phase angle for a purely resistive impedance of $600 + j0$. The R and jX scales are relative.

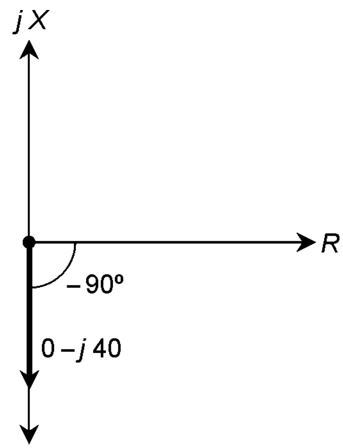
Problem 17-3

Suppose a circuit contains a resistance of 50 Ω and an inductive reactance of 50 Ω in series. What is the power factor?

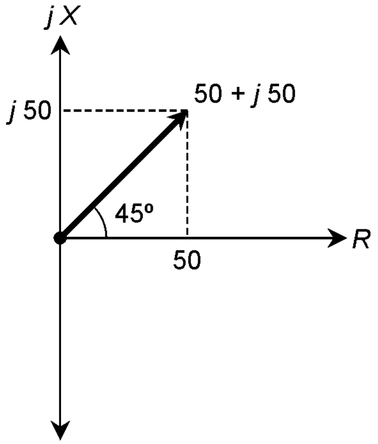
The phase angle in this case is 45° (Fig. 17-7). The resistance and reactance vectors have equal lengths and form two sides of a right triangle, with the complex impedance vector forming the hypotenuse. To determine the power factor, you can use a calculator to find cos 45° = 0.707. This means that $P_T/P_{VA} = 0.707 = 70.7\%$.

The Ratio *R/Z*

The second way to calculate the power factor is to find the ratio of the resistance R to the absolute-value impedance Z. In Fig. 17-7, this is visually apparent. A right triangle is formed by the resistance vector R (the base), the reactance vector jX (the height), and the absolute-value impedance Z (the hypotenuse). The cosine of the phase angle is equal to the ratio of the base length to the hypotenuse length; this represents R/Z.



17-6 Vector diagram showing the phase angle for a purely capacitive impedance of $0 - j40$. The R and jX scales are relative.



17-7 Vector diagram showing the phase angle for a complex impedance of 50 + *j*50. The *R* and *jX* scales are relative.

Problem 17-4

Suppose a circuit has an absolute-value impedance Z of 100 Ω, with a resistance $R = 80$ Ω. What is the power factor?

Simply find the ratio $PF = R/Z = 80/100 = 0.8 = 80\%$. Note that it doesn't matter whether the reactance in this circuit is capacitive or inductive.

Problem 17-5

Suppose a circuit has an absolute-value impedance of 50 Ω, purely resistive. What is the power factor?

Here, $R = Z = 50$ Ω. Therefore, $PF = R/Z = 50/50 = 1 = 100\%$.

Problem 17-6

Suppose a circuit has a resistance of 50 Ω and a capacitive reactance of −30 Ω in series. What is the power factor? Use the cosine method.

First, find the phase angle. Remember the formula: $\phi = \arctan (X/R)$, where X is the reactance and R is the resistance. Therefore, $\phi = \arctan (-30/50) = \arctan (-0.60) = -31°$. The power factor is the cosine of this angle; $PF = \cos (-31°) = 0.86 = 86\%$.

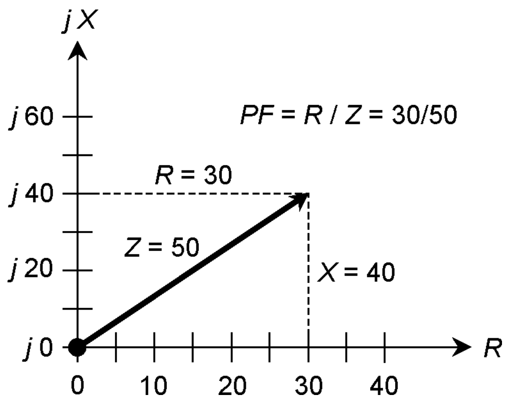
Problem 17-7

Suppose a circuit has a resistance of 30 Ω and an inductive reactance of 40 Ω. What is the power factor? Use the R/Z method.

Find the absolute-value impedance: $Z^2 = R^2 + X^2 = 30^2 + 40^2 = 900 + 1600 = 2500$. Therefore, $Z = 2500^{1/2} = 50$ Ω, so $PF = R/Z = 30/50 = 0.60 = 60\%$. This problem can be represented vectorially by a 30:40:50 right triangle, as shown in Fig. 17-8.

How Much of the Power Is True?

The preceding formulas allow you to figure out, given the resistance, reactance, and VA power, how many watts are true or real power, and how many watts are imaginary or reactive power. This is important in RF equipment, because some RF wattmeters display VA power rather than true power. When there is reactance in a circuit or system, the wattage reading is therefore exaggerated.



17-8 Illustration for Problem 17-7. (The vertical and horizontal scale increments differ; this is a common practice in graphs, often done for illustration convenience.)

Problem 17-8

Suppose a circuit has 50 Ω of resistance and 30 Ω of inductive reactance in series. A wattmeter shows 100 W, representing the VA power. What is the true power?

First, calculate the power factor. Suppose you use the phase-angle method. Then:

$$\begin{aligned}\phi &= \arctan\,(X/R)\\ &= \arctan\,(30/50) = 31°\end{aligned}$$

The power factor is the cosine of the phase angle. Thus:

$$PF = \cos 31° = 0.86 = 86\%$$

Remember that $PF = P_T/P_{VA}$. This formula can be rearranged to solve for true power:

$$\begin{aligned}P_T &= PF \times P_{VA}\\ &= 0.86 \times 100\\ &= 86\text{ W}\end{aligned}$$

Problem 17-9

Suppose a circuit has a resistance of 1000 Ω in parallel with a capacitance of 1000 pF. The frequency is 100 kHz. If a wattmeter designed to read VA power shows a reading of 88.0 W, what is the true power?

This problem is rather complicated because the components are in parallel. To begin, be sure the units are all in agreement so the formulas will work right. Convert the frequency to megahertz: f = 100 kHz = 0.100 MHz. Convert capacitance to microfarads: C = 1000 pF = 0.001000 μF. From the previous chapter, recall the formula for capacitive susceptance, and calculate it for this situation:

$$\begin{aligned}B_C &= 6.28fC\\ &= 6.28 \times 0.100 \times 0.001000\\ &= 0.000628\text{ S}\end{aligned}$$

The conductance of the resistor, G, is the reciprocal of the resistance, R, as follows:

$$\begin{aligned} G &= 1/R \\ &= 1/1000 \\ &= 0.001000 \text{ S} \end{aligned}$$

Now, use the formulas for calculating resistance and reactance in terms of conductance and susceptance in parallel circuits. First, find the resistance:

$$\begin{aligned} R &= G/(G^2 + B^2) \\ &= 0.001000/(0.001000^2 + 0.000628^2) \\ &= 0.001000/0.000001394 \\ &= 717\ \Omega \end{aligned}$$

Then, find the reactance:

$$\begin{aligned} X &= -B/(G^2 + B^2) \\ &= -0.000628/0.000001394 \\ &= -451\ \Omega \end{aligned}$$

Next, calculate the phase angle:

$$\begin{aligned} \phi &= \arctan\ (X/R) \\ &= \arctan\ (-451/717) \\ &= \arctan\ (-0.629) \\ &= -32.2^\circ \end{aligned}$$

The power factor is found from the phase angle as follows:

$$\begin{aligned} PF &= \cos\phi \\ &= \cos\ (-32.2^\circ) \\ &= 0.846 = 84.6\% \end{aligned}$$

The VA power, P_{VA}, is given as 88.0 W. Therefore:

$$\begin{aligned} P_T &= PF \times P_{VA} \\ &= 0.846 \times 88.0 \\ &= 74.4 \text{ W} \end{aligned}$$

Power Transmission

Consider how electricity gets to your home. Generators produce large voltages and currents at a power plant. The problem: getting the electricity from the plant to the homes, businesses, and other facilities that need it. This process involves the use of long wire transmission lines. Transformers are also required to step the voltages up or down. As another example, consider a radio broadcast or communications station. The transmitter produces high-frequency ac. The problem is getting the power to be radiated by the antenna, located some distance from the transmitter. This involves the use of an RF transmission line. The most common type is coaxial cable. Two-wire line is also sometimes used. At ultrahigh and microwave frequencies, another kind of transmission line, known as a *waveguide*, is often employed.

Loss: The Less, The Better!

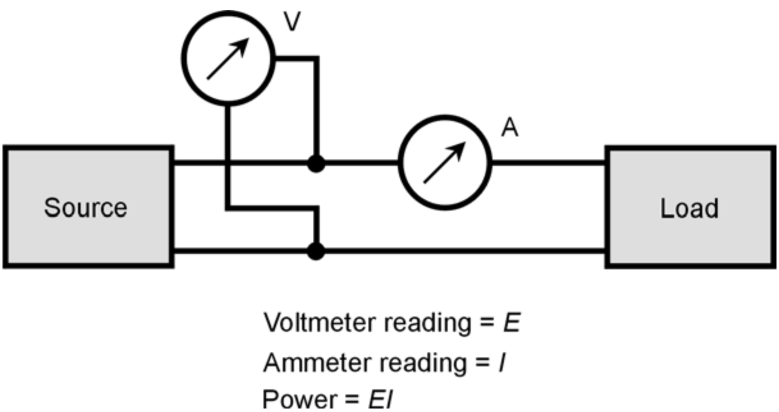
The overriding concern in any *power transmission* system is minimizing the loss. Power wastage occurs almost entirely as heat in the transmission line conductors and dielectric, and in objects near the line. Some loss can take the form of unwanted electromagnetic radiation from the line. Loss also occurs in transformers. Power loss in an electrical system is analogous to the loss of usable work produced by friction in a mechanical system. The less of it, the better!

In an ideal power transmission system, all of the power is VA power; that is, it is in the form of ac in the conductors and an alternating voltage between them. It is undesirable to have power in a transmission line or transformer exist in the form of true power, because that translates into either heat loss, or radiation loss, or both. The place for true power dissipation or radiation is in the load, such as electrical appliances or radio antennas.

Power Measurement in a Transmission Line

In an ac transmission line, power is measured by placing an ac voltmeter between the conductors, and an ac ammeter in series with one of the conductors (Fig. 17-9). Then the power P (in watts) is equal to the product of the rms voltage E (in volts) and the rms current I (in amperes). This technique can be used in any transmission line. But this is not necessarily an indication of the true power dissipated by the load at the end of the line.

Recall that any transmission line has a *characteristic impedance*. This value, Z_o, depends on the diameters of the line conductors, the spacing between the conductors, and the type of dielectric material that separates the conductors. If the load is a pure resistance R containing no reactance, and if $R = Z_o$, then the power indicated by the voltmeter/ammeter scheme will be the same as the true power dissipated by the load—provided that the voltmeter and ammeter are placed at the load end of the transmission line.



17-9 Power measurement in a transmission line. Ideally, the voltage and the current should be measured at the same physical point on the line.

If the load is a pure resistance but it differs from the characteristic impedance of the line, then the voltmeter and ammeter will not give an indication of the true power. Also, if there is any reactance in the load, the voltmeter/ammeter method will not be accurate, even if the resistive component happens to be the same as the characteristic impedance of the line. The physics of this is rather complicated, and we won't get into the details here. But you should remember that it is optimum for the impedance of a load to be a pure resistance R, such that $R = Z_o$. When this is not the case, an *impedance mismatch* is said to exist.

Small impedance mismatches can often be tolerated in power transmission systems. But this is not always the case. In very high frequency (VHF), ultrahigh frequency (UHF), and microwave radio transmitting systems, even a small impedance mismatch between the load and the line can cause excessive power losses in the line. An impedance mismatch can usually be corrected by means of a *matching transformer* between a transmission line and the load, and/or the deliberate addition of reactance at the load end of the line to cancel out any existing load reactance.

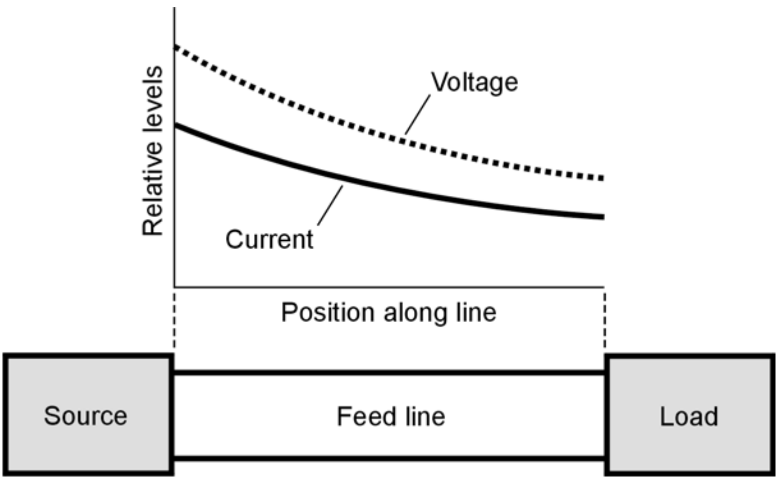
Loss in a Mismatched Line

When a transmission line is terminated in a resistance $R = Z_o$, then the current and the voltage are constant all along the line, provided the line has no loss. The ratio of the voltage to the current, E/I, is equal to R and also equal to Z_o. But this is an idealized case. No line is completely *lossless.*

In a real-world transmission line, the current and voltage gradually decrease as a signal makes its way from the source to the load. But if the load is a pure resistance equal to the characteristic impedance of the line, the current and voltage remain in the same ratio at all points along the line (Fig. 17-10).

Standing Waves

If the load is not perfectly matched to the line, the current and voltage vary in a complicated way along the length of the line. In some places, the current is high; in other places it is low. The max-



17-10 In a matched line, the ratio of the voltage to the current (E/I) is constant everywhere along the line, although the actual values of E and I decrease with increasing distance from the source.

ima and minima are called *loops* and *nodes*, respectively. At a *current loop*, the voltage is minimum (a *voltage node*), and at a *current node*, the voltage is maximum (a *voltage loop*). The current and voltage loops and nodes along a mismatched transmission line, if graphed as functions of the position on the line, form wavelike patterns that remain fixed over time. They just stand there. For this reason, they are called *standing waves.*

Standing-Wave Loss

At current loops, the loss in line conductors reaches a maximum. At voltage loops, the loss in the dielectric reaches a maximum. At current nodes, the loss in the conductors reaches a minimum. At voltage nodes, the loss in the dielectric reaches a minimum. It is tempting to suppose that everything would average out here, but it doesn't work that way! Overall, in a mismatched line, the line losses are greater than they are in a perfectly matched line. This extra line loss increases as the mismatch gets worse.

Transmission-line mismatch loss, also called *standing-wave loss*, occurs in the form of heat dissipation. It is true power. Any true power that goes into heating up a transmission line is wasted, because it cannot be dissipated in the load.

The greater the mismatch, the more severe the standing-wave loss becomes. The more loss a line has to begin with (that is, when it is perfectly matched), the more loss is caused by a given amount of mismatch. Standing-wave loss also increases as the frequency increases, if all other factors are held constant. This loss is the most significant, and the most harmful, in long lengths of transmission line, especially in RF practice at VHF, UHF, and microwave frequencies.

Line Overheating

A severe mismatch between the load and the transmission line can cause another problem: physical damage to, or destruction of, the line!

A feed line might be able to handle a kilowatt (1 kW) of power when it is perfectly matched. But if a severe mismatch exists and you try to feed 1 kW into the line, the extra current at the current loops can heat the conductors to the point where the dielectric material melts and the line shorts out. It is also possible for the voltage at the voltage loops to cause arcing between the line conductors. This perforates and/or burns the dielectric, ruining the line.

When an RF transmission line must be used with a mismatch, *derating functions* are required to determine how much power the line can safely handle. Manufacturers of prefabricated lines such as coaxial cable can supply you with this information.

Resonance

One of the most important phenomena in ac circuits, especially in RF engineering, is the property of *resonance*. This is a condition that occurs when capacitive and inductive reactance cancel each other out.

Series Resonance

Recall that capacitive reactance, X_C, and inductive reactance, X_L, can be equal in magnitude, although they are always opposite in effect. In any circuit containing an inductance and capacitance, there exists a frequency at which $X_L = -X_C$. This condition constitutes resonance. In a simple *LC* circuit, there is only one such frequency. But in some circuits involving transmission lines or antennas,

17-11 A series *RLC* circuit.

there can be many such frequencies. The lowest frequency at which resonance occurs is called the *resonant frequency*, symbolized f_o.

Refer to the schematic diagram of Fig. 17-11. You should recognize this as a series *RLC* circuit. At some particular frequency, $X_L = -X_C$. This is inevitable if L and C are finite and nonzero. This frequency is f_o for the circuit. At f_o, the effects of capacitive reactance and inductive reactance cancel out. The result is that the circuit appears as a pure resistance, with a value that is theoretically equal to R.

If $R = 0$, that is, if the resistor is a short circuit, then the circuit is called a *series LC circuit*, and the impedance at resonance will be theoretically $0 + j0$. The circuit will offer no opposition to the flow of alternating current at the frequency f_o. This condition is *series resonance.* In a practical series *LC* circuit, there is always a little bit of loss in the coil and capacitor, so the real part of the complex impedance is not exactly equal to 0 (although it can be extremely small).

Parallel Resonance

Refer to the circuit diagram of Fig. 17-12. This is a *parallel RLC circuit.* Remember that, in this sort of situation, the resistance R should be thought of as a conductance G, with $G = 1/R$. Then the circuit can be called a *parallel GLC circuit.*

At some particular frequency f_o, the inductive susceptance B_L will exactly cancel the capacitive susceptance B_C; that is, $B_L = -B_C$. This is inevitable for some frequency f_o, as long as the circuit contains finite, nonzero inductance and finite, nonzero capacitance. At the frequency f_o, the susceptances cancel each other out, leaving theoretically zero susceptance. The admittance through the circuit is then very nearly equal to the conductance, G, of the resistor.

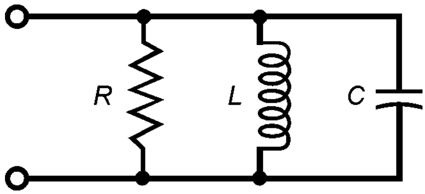
If the circuit contains no resistor, but only a coil and capacitor, it is called a *parallel LC circuit*, and the admittance at resonance will be theoretically $0 + j0$. That means the circuit will offer great opposition to alternating current at f_o, and the complex impedance will theoretically be infinite! This condition is *parallel resonance.* In a practical parallel *LC* circuit, there is always a little bit of loss in the coil and capacitor, so the real part of the complex impedance is not infinite (although it can be extremely large).

Calculating Resonant Frequency

The formula for calculating resonant frequency f_o, in terms of the inductance L in henrys and the capacitance C in farads, is as follows:

$$f_o = 1/[2\pi(LC)^{1/2}]$$

17-12 A parallel *RLC* circuit.



Considering $\pi = 3.14$ to three significant figures, this formula can be simplified to:

$$f_o = 0.159/(LC)^{1/2}$$

The $\frac{1}{2}$ power of a quantity represents the positive square root of that quantity. The preceding formulas are valid for series-resonant and parallel-resonant *RLC* circuits.

The formula will also work if you want to find f_o in megahertz (MHz) when L is given in microhenrys (μH) and C is in microfarads (μF). These values are far more common than hertz, henrys, and farads in electronic circuits. Just remember that millions of hertz go with million*th*s of henrys, and with million*th*s of farads.

The Effects of *R* and *G*

Interestingly, the value of R or G does not affect the resonant frequency in either type of circuit. But these quantities are significant, nevertheless! The presence of nonzero resistance in a *series-resonant circuit*, or nonzero conductance in a *parallel-resonant circuit*, makes the resonant frequency less well-defined. Engineers say that the *resonant frequency response* becomes "more broad" or "less sharp."

In a series circuit, the resonant frequency response becomes more broad as the *resistance* increases. In a parallel circuit, the resonant frequency response becomes more broad as the *conductance* increases. The sharpest possible responses occur when $R = 0$ in a series circuit, and when $G = 0$ (that is, $R = \infty$) in a parallel circuit.

Problem 17-10

Find the resonant frequency of a series circuit with an inductance of 100 μH and a capacitance of 100 pF.

First, convert the capacitance to microfarads: 100 pF = 0.000100 μF. Then find the product $LC = 100 \times 0.000100 = 0.0100$. Take the square root of this, getting 0.100. Finally, divide 0.159 by 0.100, getting $f_o = 1.59$ MHz.

Problem 17-11

Find the resonant frequency of a parallel circuit consisting of a 33-μH coil and a 47-pF capacitor.

Again, convert the capacitance to microfarads: 47 pF = 0.000047 μF. Then find the product $LC = 33 \times 0.000047 = 0.00155$. Take the square root of this, getting 0.0394. Finally, divide 0.159 by 0.0394, getting $f_o = 4.04$ MHz.

Problem 17-12

Suppose you want to design a circuit so that it has $f_o = 9.00$ MHz. You have a 33-pF fixed capacitor available. What size coil will be needed to get the desired resonant frequency?

Use the formula for the resonant frequency, and plug in the values. This will allow you to use simple arithmetic to solve for L. Convert the capacitance to microfarads: 33 pF = 0.000033 μF. Then calculate as follows:

$$\begin{aligned} f_o &= 0.159/(LC)^{1/2} \\ 9.00 &= 0.159/(L \times 0.000033)^{1/2} \\ 9.00^2 &= 0.159^2/(0.000033 \times L) \\ 81.0 &= 0.0253/(0.000033 \times L) \end{aligned}$$

$$
\begin{aligned}
81.0 \times 0.000033 \times L &= 0.0253 \\
0.00267 \times L &= 0.0253 \\
L &= 0.0253/0.00267 \\
&= 9.48\ \mu H
\end{aligned}
$$

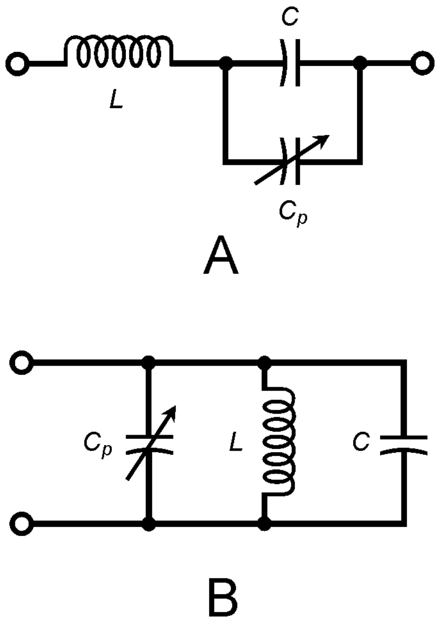
Problem 17-13

Suppose a circuit must be designed to have $f_o = 455$ kHz. A coil of 100 μH is available. What size capacitor is needed?

Convert the frequency to megahertz: 455 kHz = 0.455 MHz. Then the calculation proceeds in the same way as with the preceding problem:

$$
\begin{aligned}
f_o &= 0.159/(LC)^{1/2} \\
0.455 &= 0.159/(100 \times C)^{1/2} \\
0.455^2 &= 0.159^2/(100 \times C) \\
0.207 &= 0.0253/(100 \times C) \\
0.207 \times 100 \times C &= 0.0253 \\
20.7 \times C &= 0.0253 \\
C &= 0.0253/20.7 \\
&= 0.00122\ \mu\text{F} \\
&= 1220\ \text{pF}
\end{aligned}
$$

In practical circuits, variable inductors and/or variable capacitors are often placed in tuned circuits, so that small errors in the frequency can be compensated for. The most common approach is to design the circuit for a frequency slightly higher than f_o, and to use a *padder capacitor* in parallel with the main capacitor (Fig. 17-13).



17-13 Padding capacitors (C_p) allow limited adjustment of the resonant frequency in a series *LC* circuit (as shown at A), or in a parallel *LC* circuit (as shown at B).

Resonant Devices

Resonant circuits often consist of coils and capacitors in series or parallel, but there are other kinds of hardware that exhibit resonance. Some of these are as follows.

Piezoelectric Crystals

Pieces of quartz, when cut into thin wafers and subjected to voltages, will vibrate at high frequencies. Because of the physical dimensions of such a *piezoelectric crystal*, these vibrations occur at a precise frequency f_o, and also at whole-number multiples of f_o. These multiples, $2f_o$, $3f_o$, $4f_o$, and so on, are called *harmonic frequencies* or simply *harmonics.* The frequency f_o is called the *fundamental frequency* or simply the *fundamental.* The fundamental, f_o, is defined as the lowest frequency at which resonance occurs. Quartz crystals can be made to act like *LC* circuits in electronic devices. A crystal exhibits an impedance that varies with frequency. The reactance is zero at f_o and the harmonic frequencies.

Cavities

Lengths of metal tubing, cut to specific dimensions, exhibit resonance at very high, ultrahigh, and microwave radio frequencies. They work in much the same way as musical instruments resonate with sound waves. But the waves are electromagnetic, rather than acoustic. Such *cavities*, also called cavity *resonators*, have reasonable physical dimensions at frequencies above about 150 MHz. Below this frequency, a cavity can be made to work, but it is long and unwieldy. Like crystals, cavities resonate at a fundamental frequency f_o, and also at harmonic frequencies.

Sections of Transmission Line

When a transmission line is cut to ¼ wavelength, or to any whole-number multiple of this, it behaves as a resonant circuit. The most common length for a transmission-line resonator is a ¼ wavelength. Such a piece of transmission line is called a *quarter-wave section.*

When a quarter-wave section is short-circuited at the far end, it acts like a parallel-resonant *LC* circuit, and has a high resistive impedance at the resonant frequency f_o. When it is open at the far end, it acts as a series-resonant *LC* circuit, and has a low resistive impedance at f_o. In effect, a quarter-wave section converts an ac short circuit into an ac open circuit and vice versa, at a specific frequency f_o.

The length of a quarter-wave section depends on the desired f_o. It also depends on how fast the electromagnetic energy travels along the line. This speed is specified in terms of a *velocity factor*, abbreviated *v.* The value of v is given as a fraction of the speed of light. Typical transmission lines have velocity factors ranging from about 0.66 to 0.95 (or 66 percent to 95 percent). This factor is provided by the manufacturers of prefabricated lines such as coaxial cable.

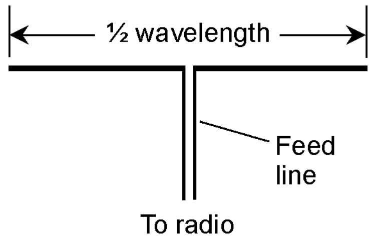
If the frequency in megahertz is f_o and the velocity factor of a line is v, then the length L_{ft} of a quarter-wave section of transmission line, in feet, is given by this formula:

$$L_{ft} = 246v/f_o$$

The length L_m in meters is given by this:

$$L_m = 75.0v/f_o$$

We use L here to stand for "length," not "inductance"!



17-14 The half-wave, center-fed dipole is a simple and efficient antenna.

Antennas

Many types of antennas exhibit resonant properties. The simplest type of resonant antenna, and the only kind that will be mentioned here, is the center-fed, half-wavelength *dipole antenna* (Fig. 17-14).

The length L_{ft}, in feet, for a dipole antenna at a frequency of f_o, in megahertz, is given by the following formula:

$$L_{ft} = 468/f_o$$

This takes into account the fact that electromagnetic fields travel along a wire at about 95 percent of the speed of light. A straight, thin wire in free space has a velocity factor of approximately 0.95.

If the length of the half-wave dipole is specified in meters as L_m, then:

$$L_m = 143/f_o$$

A half-wave dipole has a purely resistive impedance of about 73 Ω at its fundamental frequency f_o. But this type of antenna is also resonant at all harmonics of f_o. The dipole is a full wavelength long at $2f_o$; it is $3/2$ wavelength long at $3f_o$; it is two full wavelengths long at $4f_o$, and so on.

Radiation Resistance

At f_o and all of the odd harmonics, the antenna behaves like a series-resonant *RLC* circuit with a fairly low resistance. At all even harmonics, the antenna acts like a parallel-resonant *RLC* circuit with a high resistance. Does this confuse you? There's no resistor in Fig. 17-14! Where, you ask, does the resistance come from in the half-wave dipole? The answer to this is rather esoteric, and it brings to light an interesting property that all antennas have. It is called *radiation resistance*, and is a crucial factor in the design and construction of all RF antenna systems.

When electromagnetic energy is fed into an antenna, power is radiated into space in the form of radio waves. This is a manifestation of true power, just as the dissipation of power in a pure resistance is a manifestation of true power. Although there is no physical resistor in Fig. 17-14, the radiation of radio waves is like power dissipation in a pure resistance. In fact, if a half-wave dipole antenna were replaced with a 73-Ω nonreactive resistor that could dissipate enough power without burning out, a radio transmitter connected to the opposite end of the line wouldn't know the difference. (But a receiver would!)

Problem 17-14

How many feet long is a quarter-wave section of transmission line at 7.05 MHz, if the velocity factor is 0.800?

Just use the formula:

$$
\begin{aligned}
L_{\text{ft}} &= 246v/f_o \\
&= (246 \times 0.800)/7.05 \\
&= 197/7.05 \\
&= 27.9 \text{ ft}
\end{aligned}
$$

Quiz

Refer to the text in this chapter if necessary. A good source is 18 or more correct. Answers are in the back of the book.

1. The power in a pure reactance is
   (a) radiated.
   (b) true.
   (c) imaginary.
   (d) apparent.

2. Which of the following is *not* an example of true power?
   (a) Power in the form of heat, produced by dc flowing through a resistor
   (b) Power in the form of electromagnetic fields, radiated from a radio antenna
   (c) The product of the rms ac through a capacitor and the rms voltage across it
   (d) Power in the form of heat, produced by losses in an RF transmission line

3. Suppose the apparent power in a circuit is 100 W, and the imaginary power is 40 W. What is the true power?
   (a) 92 W
   (b) 100 W
   (c) 140 W
   (d) It is impossible to determine from this information.

4. Power factor is equal to
   (a) apparent power divided by true power.
   (b) imaginary power divided by apparent power.
   (c) imaginary power divided by true power.
   (d) true power divided by apparent power.

5. Suppose a circuit has a resistance of 300 Ω and an inductance of 13.5 μH in series, and is operated at 10.0 MHz. What is the power factor?
   (a) 0.334
   (b) 0.999
   (c) 0.595
   (d) It cannot be determined from the information given.

6. Suppose a series circuit has $Z = 88.4\ \Omega$, with $R = 50.0\ \Omega$. What is the power factor, expressed as a percentage?
   (a) 99.9 percent
   (b) 56.6 percent
   (c) 60.5 percent
   (d) 29.5 percent

7. Suppose a series circuit has $R = 53.5\ \Omega$, with $X = 75.5\ \Omega$. What is the power factor, expressed as a percentage?
   (a) 70.9 percent
   (b) 81.6 percent
   (c) 57.8 percent
   (d) 63.2 percent

8. The phase angle in an ac circuit is equal to
   (a) arctan (Z/R).
   (b) arctan (R/Z).
   (c) arctan (R/X).
   (d) arctan (X/R).

9. Suppose an ac ammeter and an ac voltmeter indicate that there are 220 W of VA power in a circuit that consists of a resistance of 50 Ω in series with a capacitive reactance of −20 Ω. What is the true power?
   (a) 237 W
   (b) 204 W
   (c) 88.0 W
   (d) 81.6 W

10. Suppose an ac ammeter and an ac voltmeter indicate that there are 57 W of VA power in a circuit. The resistance is known to be 50 Ω, and the true power is known to be 40 W. What is the absolute-value impedance?
    (a) 50 Ω
    (b) 57 Ω
    (c) 71 Ω
    (d) It is impossible to determine on the basis of this data.

11. Which of the following should be minimized in an RF transmission line?
    (a) The load impedance
    (b) The load resistance
    (c) The line loss
    (d) The transmitter power

12. Which of the following does *not* increase the loss in a transmission line?
    (a) Reducing the power output of the source
    (b) Increasing the degree of mismatch between the line and the load
    (c) Reducing the diameter of the line conductors
    (d) Raising the frequency

13. Which of the following is a significant problem that standing waves can cause in an RF transmission line?
    (a) Line overheating
    (b) Excessive power loss
    (c) Inaccuracy in power measurement
    (d) All of the above

14. Suppose a coil and capacitor are in series. The inductance is 88 mH and the capacitance is 1000 pF. What is the resonant frequency?
    (a) 17 kHz
    (b) 540 Hz
    (c) 17 MHz
    (d) 540 kHz

15. Suppose a coil and capacitor are in parallel, with $L = 10.0$ μH and $C = 10$ pF. What is f_o?
    (a) 15.9 kHz
    (b) 5.04 MHz
    (c) 15.9 MHz
    (d) 50.4 MHz

16. Suppose you want to build a series-resonant circuit with $f_o = 14.1$ MHz. A coil of 13.5 μH is available. How much capacitance is needed?
    (a) 0.945 μF
    (b) 9.45 pF
    (c) 94.5 pF
    (d) 945 pF

17. Suppose you want to build a parallel-resonant circuit with $f_o = 21.3$ MHz. A capacitor of 22.0 pF is available. How much inductance is needed?
    (a) 2.54 mH
    (b) 254 μH
    (c) 25.4 μH
    (d) 2.54 μH

18. A ¼-wave section of transmission line is cut for use at 21.1 MHz. The line has a velocity factor of 0.800. What is its physical length in meters?
    (a) 11.1 m
    (b) 3.55 m

(c) 8.87 m
(d) 2.84 m

19. What is the fourth harmonic of 800 kHz?
(a) 200 kHz
(b) 400 kHz
(c) 3.20 MHz
(d) 4.00 MHz

20. Suppose you want to build a $\frac{1}{2}$-wave dipole antenna designed to have a fundamental resonant frequency of 3.60 MHz. How long should you make it, as measured from end to end in feet?
(a) 130 ft
(b) 1680 ft
(c) 39.7 ft
(d) 515 ft

18
CHAPTER

Transformers and Impedance Matching

TRANSFORMERS ARE USED TO OBTAIN THE OPTIMUM VOLTAGE FOR THE OPERATION OF A CIRCUIT OR system. Transformers can also match impedances between a circuit and a load, or between two different circuits. Transformers can be used to provide dc isolation between electronic circuits while letting ac pass. Another application is to mate balanced and unbalanced circuits, feed systems, and loads.

Principle of the Transformer

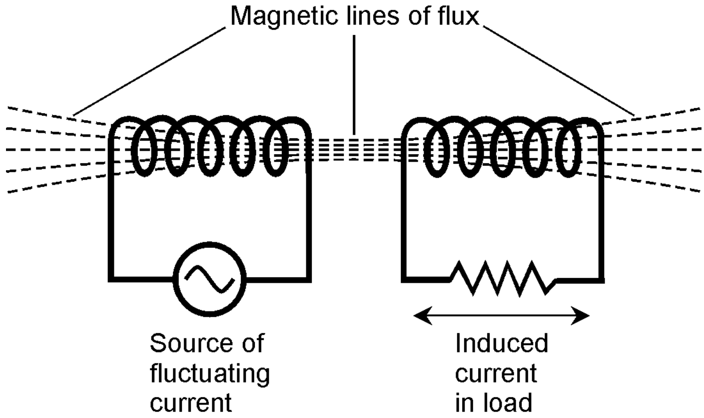
When two wires are near each other and one of them carries a fluctuating current, a fluctuating current is induced in the other wire. This effect is known as *electromagnetic induction.* All ac transformers work according to the principle of electromagnetic induction. If the first wire carries sine-wave ac of a certain frequency, then the *induced current* is sine-wave ac of the same frequency in the second wire.

The closer the two wires are to each other, the greater is the induced current, for a given current in the first wire. If the wires are wound into coils and placed along a common axis (Fig. 18-1), the induced current will be greater than if the wires are straight and parallel. Even more *coupling*, or efficiency of induced-current transfer, is obtained if the two coils are wound one atop the other.

Primary and Secondary

The two windings, along with the core on which they are wound, constitute a transformer. The first coil is called the *primary winding*, and the second coil is known as the *secondary winding.* These are often spoken of simply as the *primary* and the *secondary.* The induced current in the secondary creates a voltage between its end terminals. In a *step-down* transformer, the secondary voltage is less than the primary voltage. In a *step-up* transformer, the secondary voltage is greater than the primary voltage. The primary voltage is abbreviated E_{pri}, and the secondary voltage is abbreviated E_{sec}. Unless otherwise stated, effective (rms) voltages are always specified.

The windings of a transformer have inductance, because they are coils. The required inductances of the primary and secondary depend on the frequency of operation, and also on the resistive part of the impedance in the circuit. As the frequency increases, the needed inductance decreases. At high resistive impedances, more inductance is generally needed than at low resistive impedances.



18-1 Magnetic lines of flux between two aligned coils of wire when one of the coils carries fluctuating or alternating current.

Turns Ratio

The *primary-to-secondary turns ratio* in a transformer is the ratio of the number of turns in the primary, T_{pri}, to the number of turns in the secondary, T_{sec}. This ratio is written T_{pri}:T_{sec} or T_{pri}/T_{sec}. In a transformer with excellent primary-to-secondary coupling, the following relationship always holds:

$$E_{pri}/E_{sec} = T_{pri}/T_{sec}$$

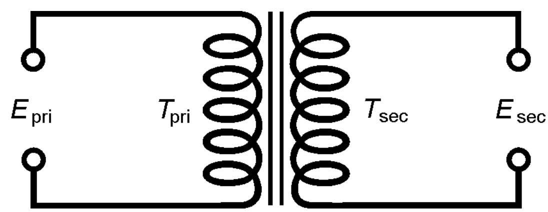
That is, the primary-to-secondary voltage ratio is always equal to the primary-to-secondary turns ratio (Fig. 18-2).

Problem 18-1

Suppose a transformer has a primary-to-secondary turns ratio of exactly 9:1. The ac voltage at the primary is 117 V rms. Is this a step-up transformer or a step-down transformer? What is the voltage across the secondary?

This is a step-down transformer. Simply plug in the numbers in the preceding equation and solve for E_{sec}, as follows:

$$\begin{aligned} E_{pri}/E_{sec} &= T_{pri}/T_{sec} \\ 117/E_{sec} &= 9.00 \\ 1/E_{sec} &= 9.00/117 \\ E_{sec} &= 117/9.00 \\ &= 13.0 \text{ V rms} \end{aligned}$$



18-2 The primary voltage (E_{pri}) and secondary voltage (E_{sec}) in a transformer depend on the number of turns in the primary winding (T_{pri}) versus the number of turns in the secondary winding (T_{sec}).

Problem 18-2

Consider a transformer with a primary-to-secondary turns ratio of exactly 1:9. The voltage at the primary is 121.4 V rms. Is this a step-up transformer or a step-down transformer? What is the voltage at the secondary?

This is a step-up transformer. Plug in numbers and solve for E_{sec}, as follows:

$$
\begin{aligned}
121.4/E_{sec} &= 1/9.000 \\
E_{sec}/121.4 &= 9.000 \\
E_{sec} &= 9.000 \times 121.4 \\
&= 1093 \text{ V rms}
\end{aligned}
$$

Sometimes the *secondary-to-primary turns ratio* is given, rather than the primary-to-secondary turns ratio. This is written T_{sec}/T_{pri}. In a step-down unit, T_{sec}/T_{pri} is less than 1. In a step-up unit, T_{sec}/T_{pri} is greater than 1. When you hear someone say that such-and-such a transformer has a certain "turns ratio," say 10:1, be sure of which ratio is meant, T_{pri}/T_{sec} or T_{sec}/T_{pri}! If you get it wrong, you'll have the secondary voltage wrong by a factor of the *square* of the turns ratio.

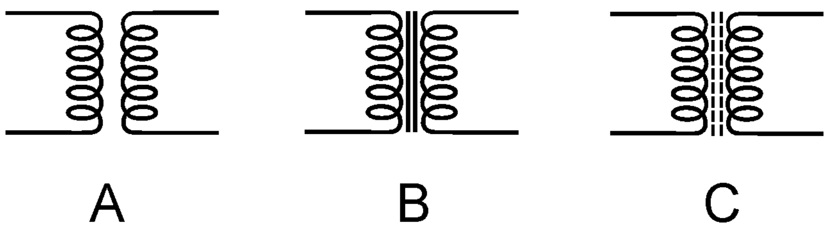
Ferromagnetic Cores

If a ferromagnetic substance such as laminated iron or powdered iron is placed within the pair of coils, the extent of coupling is increased far above that possible with an air core. But this improvement in coupling is obtained at a price. Some energy is invariably lost as heat in the core. Also, ferromagnetic cores limit the maximum frequency at which a transformer will work well.

The schematic symbol for an air-core transformer consists of two inductor symbols back-to-back (Fig. 18-3A). If a laminated iron core is used, two parallel lines are added to the schematic symbol (Fig. 18-3B). If the core is made of powdered iron, the two parallel lines are broken or dashed (Fig. 18-3C).

In transformers for 60-Hz utility ac, and also for low audio-frequency (AF) use, sheets of an alloy called *silicon steel*, glued together in layers, are often employed as transformer cores. The silicon steel is sometimes called *transformer iron.* The reason layering is used, rather than making the core from a single mass of metal, is that the magnetic fields from the coils cause currents to flow in a solid core. These *eddy currents* go in circles, heating up the core and wasting energy that would otherwise be transferred from the primary to the secondary. Eddy currents are choked off by breaking up the core into layers, so that currents cannot flow very well in circles.

A rather esoteric form of loss, called *hysteresis loss*, occurs in all ferromagnetic transformer cores, but especially laminated iron. Hysteresis is the tendency for a core material to be sluggish in accept-



18-3 Schematic symbols for transformers. At A, air core. At B, laminated iron core. At C, ferrite or powdered iron core.

ing a fluctuating magnetic field. Laminated cores exhibit high hysteresis loss above the AF range, and are therefore not good above a few kilohertz.

At frequencies up to several tens of megahertz, *powdered iron* works well for RF transformers. This material has high magnetic permeability and concentrates the flux efficiently. High permeability cores minimize the number of turns needed in the coils, and this minimizes the loss that occurs in the wires.

Geometries

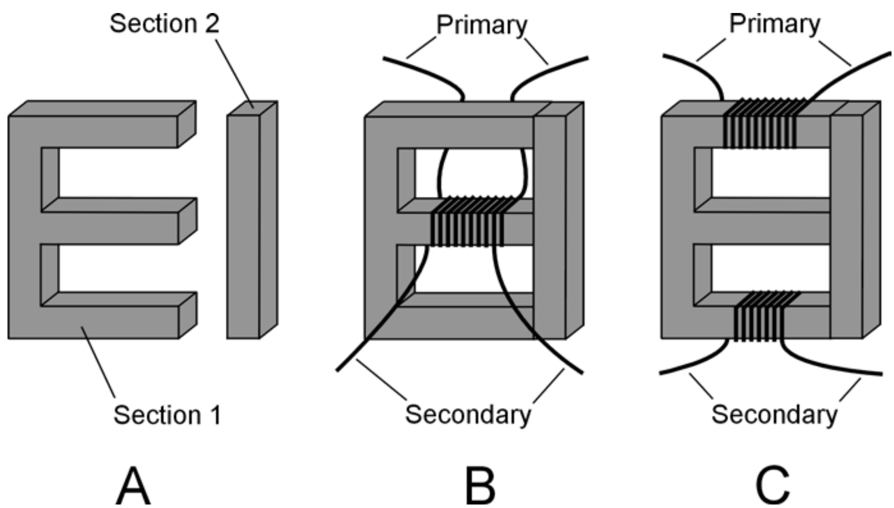
The properties of a transformer depend on the shape of its core, and on the way in which the wires are wound on it. There are several different *geometries* used with transformers.

E Core

A common core for a power transformer is the *E core*, so named because it is shaped like the capital letter E. A bar, placed at the open end of the E, completes the core assembly after the coils have been wound on the E-shaped section (Fig. 18-4A).

The primary and secondary windings can be placed on an E core in either of two ways. The simpler winding method is to put both the primary and the secondary around the middle bar of the E (Fig. 18-4B). This is called the *shell method* of transformer winding. It provides maximum coupling between the windings. However, this scheme results in considerable capacitance between the primary and the secondary. Such *interwinding capacitance* can sometimes be tolerated, but often it cannot. Another disadvantage of the shell geometry is that, when windings are placed one on top of the other, the transformer cannot handle very much voltage. High voltages cause arcing between the windings, which can destroy the insulation on the wires and lead to permanent short circuits.

Another winding method is the *core method.* In this scheme, one winding is placed at the bottom of the E section, and the other winding is placed at the top (Fig. 18-4C). The coupling occurs



18-4 At A, a utility transformer E core, showing both sections. At B, the shell winding method. At C, the core winding method.

by means of magnetic flux in the core. The interwinding capacitance is lower than it is in a shell-wound transformer because the windings are physically farther apart. Also, a core-wound transformer can handle higher voltages than a shell-wound transformer of the same physical size. Sometimes the center part of the E is left out of the core when the core winding scheme is used.

Shell-wound and core-wound transformers are almost universally employed at 60 Hz. These configurations are also common at AF.

Solenoidal Core

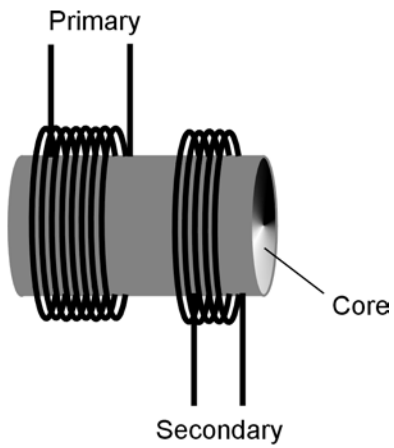
A pair of cylindrical coils, wound around a rod-shaped piece of powdered iron or ferrite, was once a common configuration for RF transformers. Sometimes this type of transformer is still seen, although it is most often used as a *loopstick antenna* in portable radio receivers and in radio direction-finding equipment. The coil windings can be placed one atop the other, or they can be separated (Fig. 18-5) to reduce the capacitance between the primary and secondary.

In a loopstick antenna, the primary serves to pick up the radio signals. The secondary winding provides an optimum impedance match to the first amplifier stage, or *front end*, of the radio receiver. The use of transformers for impedance matching is discussed later in this chapter.

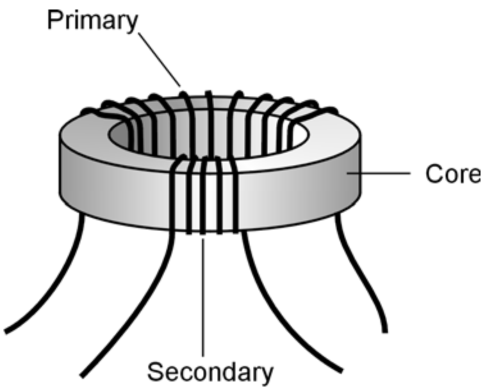
Toroidal Core

The *toroidal core* (or *toroid*) has become common for winding RF transformers. The core is a donut-shaped ring of powdered iron. The coils are wound around the donut. The complete assembly is called a *toroidal transformer.* The primary and secondary can be wound one over the other, or they can be wound over different parts of the core (Fig. 18-6). As with other transformers, when the windings are one on top of the other, there is more interwinding capacitance than when they are separated.

Toroids confine practically all the magnetic flux within the core material. This allows toroidal coils and transformers to be placed near other components without inductive interaction. Also, a toroidal coil or transformer can be mounted directly on a metal chassis, and the operation is not affected (assuming the wire is insulated or enameled).



18-5 A solenoidal-core transformer.

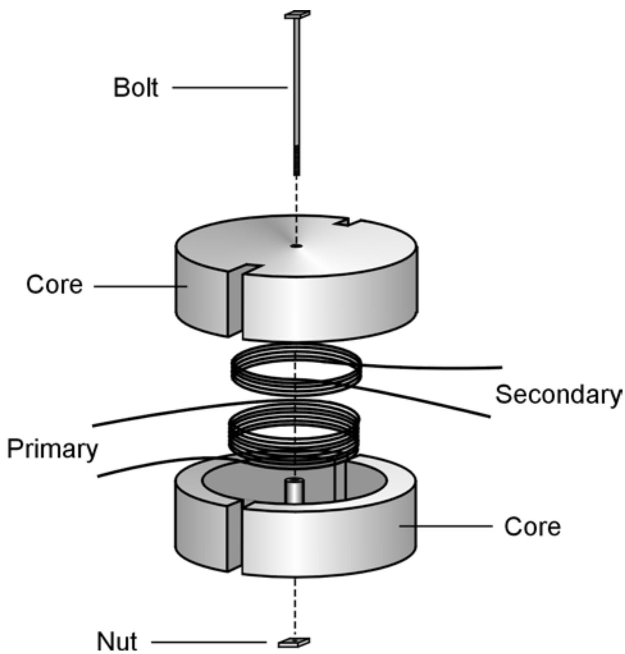


18-6 A toroidal-core transformer.

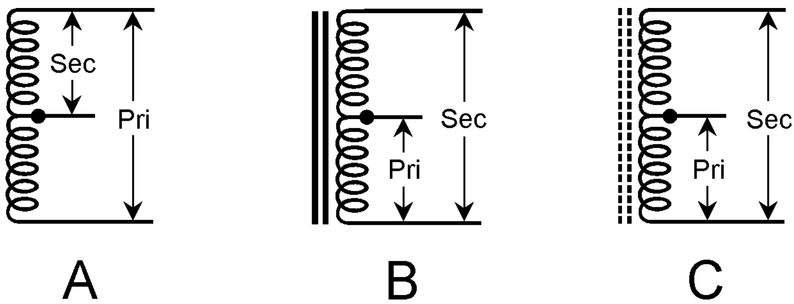
A toroidal core provides considerably more inductance per turn, for the same kind of ferromagnetic material, than a solenoidal core. It is common to see toroidal coils or transformers that have inductance values as high as 100 mH.

Pot Core

Even more inductance per turn can be obtained with a *pot core*. This is a shell of ferromagnetic material that is wrapped around a loop-shaped coil. The core is manufactured in two halves (Fig. 18-7). You wind the coil inside one of the halves, and then bolt the two together. The final core completely surrounds the loop, and the magnetic flux is confined to the core material.



18-7 Exploded view of a pot-core transformer.



18-8 Schematic symbols for autotransformers. At A, air core, step-down. At B, laminated iron core, step-up. At C, ferrite or powdered iron core, step-up.

Like the toroid, the pot core is self-shielding. There is essentially no coupling to external components. A pot core can be used to wind a single, high-inductance coil. Inductance values of more than 1 H are possible with a reasonable number of wire turns.

In a pot-core transformer, the primary and secondary must be wound next to each other. This is unavoidable because of the geometry. Therefore, the interwinding capacitance of a pot-core transformer is high. Pot cores are useful at AF and the lowest-frequency parts of the RF spectrum. They are rarely employed at high radio frequencies.

Autotransformer

In some situations, there is no need to provide dc isolation between the primary and secondary windings of a transformer. In a case of this sort, an *autotransformer* can be used. It has a single, tapped winding.

Figure 18-8 shows three autotransformer configurations. The unit shown at A has an air core, and is a step-down type. The unit at B has a laminated iron core, and is a step-up type. The unit at C has a powdered iron core, and is a step-up type.

You'll sometimes see autotransformers in radio receivers or transmitters. Autotransformers work well in impedance-matching applications, and also perform well as solenoidal loopstick antennas. Autotransformers are occasionally, but not often, used in AF applications and in 60-Hz utility wiring. In utility circuits, autotransformers can step the voltage down by a large factor, but they aren't used to step voltages up by more than a few percent.

Power Transformers

Any transformer used in the 60-Hz utility line, intended to provide a certain rms ac voltage for the operation of electrical circuits, is a *power transformer.* Power transformers exist in a vast range of physical sizes, from smaller than a tennis ball to as big as a room.

At the Generating Plant

The largest transformers are employed at the places where electricity is generated. Not surprisingly, high-energy power plants have bigger transformers that develop higher voltages than low-energy, local power plants. These transformers must be able to handle high voltages and large currents simultaneously.

When electrical energy must be sent over long distances, extremely high voltages are used. This is because, for a given amount of power ultimately dissipated by the loads, the current is lower when the voltage is higher. Lower current translates into reduced loss in the transmission line.

Recall the formula $P = EI$, where P is the power (in watts), E is the voltage (in volts), and I is the current (in amperes). If you can make the voltage 10 times larger, for a given power level, then the current is reduced to $1/10$ as much. The *ohmic losses* in the wires are proportional to the *square* of the current. Remember that $P = I^2R$, where P is the power (in watts), I is the current (in amperes), and R is the resistance (in ohms). Engineers can't do much about the wire resistance or the power consumed by the loads, but they can adjust the voltage, and thereby the current.

Suppose the voltage in a power transmission line is increased by a factor of 10, and the load at the end of the line draws constant power. This increase in the voltage reduces the current to $1/10$ of its previous value. As a result, the ohmic loss is cut to $(1/10)^2$, or $1/100$, of its previous amount. That's a major improvement in the efficiency of the transmission line, at least in terms of the loss caused by the resistance in the wires—and it is the reason why regional power plants have massive transformers capable of generating hundreds of thousands of volts.

Along the Line

Extreme voltage is good for *high-tension* power transmission, but it's certainly of no use to an average consumer. The wiring in a high-tension system must be done using precautions to prevent arcing (sparking) and short circuits. Personnel must be kept at least several meters away from the wires. Can you imagine trying to use an appliance, say a home computer, by plugging it into a 500,000-V rms electrical outlet?

Medium-voltage power lines branch out from the major lines, and step-down transformers are used at the branch points. These lines fan out to still lower-voltage lines, and step-down transformers are employed at these points, too. Each transformer must have windings heavy enough to withstand the product $P = EI$, the amount of VA power delivered to all the subscribers served by that transformer, at periods of peak demand.

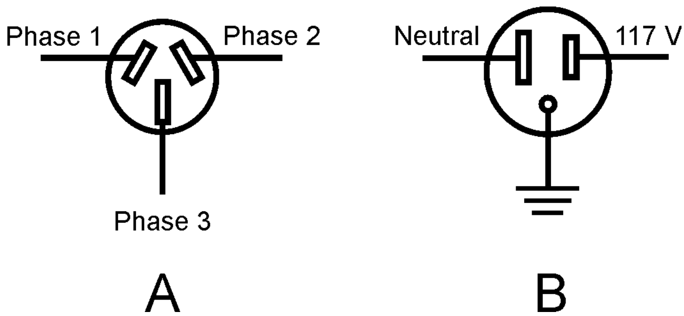
Sometimes, such as during a heat wave, the demand for electricity rises above the normal peak level. This loads down the circuit to the point that the voltage drops several percent. This is called a *brownout.* If consumption rises further still, a dangerous current load is placed on one or more intermediate power transformers. Circuit breakers in the transformers protect them from destruction by opening the circuit. Then there is a temporary *blackout.*

At individual homes and buildings, transformers step the voltage down to either 234 V rms or 117 V rms. Usually, 234-V rms electricity is provided in the form of three sine waves, called *phases*, each separated by 120°, and each appearing at one of the three slots in the outlet (Fig. 18-9A). This voltage is commonly employed with heavy appliances, such as the kitchen oven/stove (if they are electric), heating (if it is electric), and the laundry washer and dryer. A 117-V rms outlet supplies just one phase, appearing between two of the three slots in the outlet. The third opening in the outlet leads to an earth ground (Fig. 18-9B).

In Electronic Devices

The smallest power transformers are found in electronic equipment such as television sets, ham radios, and home computers. Most solid-state devices use low voltages, ranging from about 5 V up to perhaps 50 V. This equipment needs step-down power transformers in its power supplies.

Solid-state equipment usually (but not always) consumes relatively little power, so the transformers are usually not very bulky. The exception is high-powered AF or RF amplifiers, whose tran-



18-9 At A, an outlet for three-phase, 234-V rms utility ac. At B, a conventional single-phase utility outlet for 117-V rms utility ac.

sistors can demand more than 1000 W (1 kW) in some cases. At 12 V, this translates to a current demand of 90 A or more.

Television sets have cathode-ray tubes that need several hundred volts. This is derived by using a step-up transformer in the power supply. Such transformers don't have to supply a lot of current, though, so they are not very big or heavy. Another type of device that needs rather high voltage is a ham-radio amplifier with vacuum tubes. Such an amplifier requires from 2 kV to 5 kV.

Any voltage higher than about 12 V should be treated with respect. *Warning: The voltages in televisions and ham radios can present an electrocution hazard, even after the equipment has been switched off. Do not try to service such equipment unless you are trained to do so!*

At Audio Frequencies

Transformers for use at AF are similar to those employed for 60-Hz electricity. The differences are that the frequency is somewhat higher (up to 20 kHz), and that audio signals exist in a band of frequencies (20 Hz to 20 kHz) rather than at only one frequency.

Most AF transformers are constructed like miniature utility transformers. They have laminated E cores with primary and secondary windings wound around the crossbars, as shown in Fig. 18-4. Audio transformers can be either the step-up or the step-down type. However, rather than being made to produce a specific voltage, AF transformers are designed to match impedances.

Audio circuits, and in fact all electronic circuits that handle sine-wave or complex-wave signals, exhibit impedance at the input and output. The load has a certain impedance; a source has another impedance. Good audio design strives to minimize the reactance in the circuitry, so that the absolute-value impedance Z is close to the resistance R. This means that X must be zero or nearly zero. In the following discussion of impedance-matching transformers, for both AF and RF applications, assume that the reactance is zero, so the impedance is purely resistive with $Z = R + j0$.

Isolation and Impedance Matching

Transformers can provide *isolation* between electronic circuits. While there is *inductive coupling* in a transformer, there is comparatively little *capacitive coupling*. The amount of capacitive coupling can be reduced by using cores that minimize the number of wire turns needed in the windings, and by keeping the windings physically separated from each other (rather than overlapping).

Balanced and Unbalanced Loads and Lines

A *balanced load* is one whose terminals can be reversed without significantly affecting circuit behavior. A plain resistor is a good example. The two-wire antenna input in a television receiver is another example of a balanced load. A *balanced transmission line* is usually a two-wire line, such as old-fashioned *TV ribbon*, also called *twinlead.*

An *unbalanced load* is a load that must be connected a certain way. Switching its leads will result in improper circuit operation. In this sense, an unbalanced load is a little like a polarized component such as a battery or capacitor. Many wireless antennas are of this type. Usually, unbalanced sources and loads have one side connected to ground. The coaxial input of a television receiver is unbalanced; the shield (braid) of the cable is grounded. An *unbalanced transmission line* is usually a coaxial line, such as you find in a cable television system.

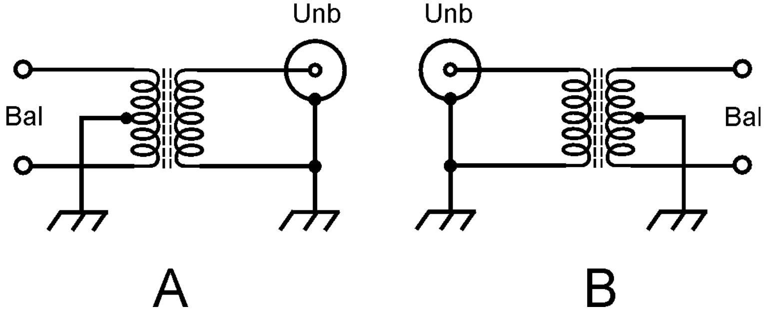
Normally, you cannot connect an unbalanced line to a balanced load, or a balanced line to an unbalanced load, and expect good performance. But a transformer can allow for mating between these two types of systems. In Fig. 18-10A, a *balanced-to-unbalanced transformer* is shown. Note that the balanced side is center-tapped, and the tap is grounded. In Fig. 18-10B, an *unbalanced-to-balanced transformer* is illustrated. Again, the balanced side has a grounded center tap.

The turns ratio of a balanced-to-unbalanced transformer (also called a *balun*) or an unbalanced-to-balanced transformer (also known as an *unbal*) can be 1:1, but this need not be the case, and often it is not. If the impedances of the balanced and unbalanced parts of the systems are the same, then a 1:1 turns ratio is ideal. But if the impedances differ, the turns ratio should be such that the impedances are matched. Shortly, we'll see how the turns ratio of a transformer can be manipulated to transform one purely resistive impedance into another.

Transformer Coupling

Transformers are sometimes used between amplifier stages in electronic equipment where a large *amplification factor* is needed. There are other methods of *coupling* from one amplifier stage to another, but transformers offer some advantages, especially in RF receivers and transmitters.

Part of the problem in getting a radio to work is that the amplifiers must operate in a stable manner. If there is too much feedback, a series of amplifiers will oscillate, and this will severely degrade the performance of the radio. Transformers that minimize the capacitance between the amplifier stages, while still transferring the desired signals, can help to prevent this oscillation.



18-10 At A, a balanced-to-unbalanced transformer. At B, an unbalanced-to-balanced transformer.

Impedance Transfer Ratio

In RF and AF systems, transformers are employed to match impedances. Thus, you will sometimes hear or read about an *impedance step-up transformer* or an *impedance step-down transformer.*

The *impedance transfer ratio* of a transformer varies according to the square of the turns ratio, and also according to the square of the voltage-transfer ratio. If the primary (source) and secondary (load) impedances are purely resistive and are denoted Z_{pri} and Z_{sec}, then the following relations hold:

$$Z_{pri}/Z_{sec} = (T_{pri}/T_{sec})^2$$
$$Z_{pri}/Z_{sec} = (E_{pri}/E_{sec})^2$$

The inverses of these formulas, in which the turns ratio or voltage-transfer ratio are expressed in terms of the impedance-transfer ratio, are:

$$T_{pri}/T_{sec} = (Z_{pri}/Z_{sec})^{1/2}$$
$$E_{pri}/E_{sec} = (Z_{pri}/Z_{sec})^{1/2}$$

Problem 18-3

Consider a situation in which a transformer is needed to match an input impedance of 50.0 Ω, purely resistive, to an output impedance of 300 Ω, also purely resistive. What is the required turns ratio T_{pri}/T_{sec}?

The required transformer will have a step-up impedance ratio of $Z_{pri}/Z_{sec} = 50.0/300 = 1/6.00$. From the preceding formulas:

$$\begin{aligned} T_{pri}/T_{sec} &= (Z_{pri}/Z_{sec})^{1/2} \\ &= (1/6.00)^{1/2} \\ &= 0.16667^{1/2} \\ &= 0.408 \\ &= 1/2.45 \end{aligned}$$

Problem 18-4

Suppose a transformer has a primary-to-secondary turns ratio of 4.00:1. The load, connected to the transformer output, is a pure resistance of 37.5 Ω. What is the impedance at the primary?

The impedance-transfer ratio is equal to the square of the turns ratio. Therefore:

$$\begin{aligned} Z_{pri}/Z_{sec} &= (T_{pri}/T_{sec})^2 \\ &= (4.00/1)^2 \\ &= 4.00^2 \\ &= 16.0 \end{aligned}$$

We know that the secondary impedance, Z_{sec} is 37.5 Ω. Thus:

$$\begin{aligned} Z_{pri} &= 16.0 \times Z_{sec} \\ &= 16.0 \times 37.5 \\ &= 600\ \Omega \end{aligned}$$

Radio-Frequency Transformers

In radio receivers and transmitters, transformers can be categorized generally by the method of construction used. Some have primary and secondary windings, just like utility and audio units. Others employ transmission-line sections. These are the two most common types of transformer found at radio frequencies.

Wire-Wound Types

In wire-wound RF transformers, powdered-iron cores can be used up to quite high frequencies. Toroidal cores are common, because they are self-shielding (all of the magnetic flux is confined within the core material). The number of turns depends on the frequency, and also on the permeability of the core.

In high-power applications, air-core coils are often preferred. Although air has low permeability, it has negligible hysteresis loss, and will not heat up or fracture as powdered-iron cores sometimes do. The disadvantage of air-core coils is that some of the magnetic flux extends outside of the coil. This affects the performance of the transformer when it must be placed in a cramped space, such as in a transmitter final-amplifier compartment.

A major advantage of coil-type transformers, especially when they are wound on toroidal cores, is that they can be made to work over a wide band of frequencies, such as from 3.5 MHz to 30 MHz. These are called *broadband transformers.*

Transmission-Line Types

As you recall, any transmission line has a characteristic impedance, or Z_o, that depends on the line construction. This property is sometimes used to make impedance transformers out of coaxial or parallel-wire line.

Transmission-line transformers are always made from quarter-wave sections. From the previous chapter, remember the formula for the length of a quarter-wave section:

$$L_{ft} = 246v/f_o$$

where L_{ft} is the length of the section in feet, v is the velocity factor expressed as a fraction, and f_o is the frequency of operation in megahertz. If the length L_m is specified in meters, then:

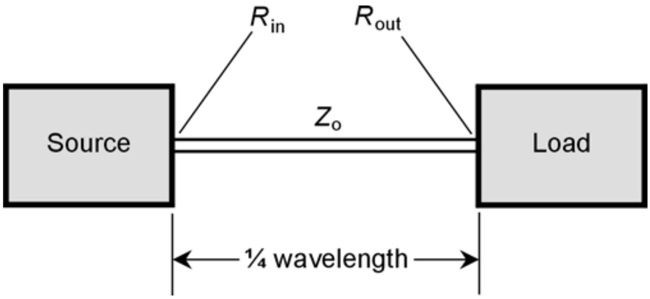
$$L_m = 75v/f_o$$

Suppose that a quarter-wave section of line, with characteristic impedance Z_o, is terminated in a purely resistive impedance R_{out}. Then the impedance that appears at the input end of the line, R_{in}, is also a pure resistance, and the following relations hold:

$$Z_o^2 = R_{in}R_{out}$$
$$Z_o = (R_{in}R_{out})^{1/2}$$

This is illustrated in Fig. 18-11. The first of the preceding formulas can be rearranged to solve for R_{in} in terms of R_{out}, or vice versa:

$$R_{in} = Z_o^2/R_{out}$$
$$R_{out} = Z_o^2/R_{in}$$



18-11 A quarter-wave matching section of transmission line. The input impedance is R_{in}, the output impedance is R_{out}, and the characteristic impedance of the line is Z_o.

These equations are valid at the frequency f_o for which the line length measures ¼ wavelength. Sometimes, the word "wavelength" is replaced by the lowercase Greek letter lambda (λ), so you will occasionally see the length of a quarter-wave section denoted as (¼)λ or 0.25λ.

Neglecting line losses, the preceding relations hold at all *odd harmonics* of f_o, that is, at $3f_o$, $5f_o$, $7f_o$, and so on. At other frequencies, a quarter-wave section of line does not act as a transformer. Instead, it behaves in a complex manner that is beyond the scope of this discussion.

Quarter-wave transmission-line transformers are most often used in antenna systems, especially at the higher frequencies, where their dimensions become practical. A quarter-wave matching section should be made using unbalanced line if the load is unbalanced, and balanced line if the load is balanced.

A disadvantage of quarter-wave sections is the fact that they work only at specific frequencies. But this is often offset by the ease with which they are constructed, if radio equipment is to be used at only one frequency, or at odd-harmonic frequencies.

Problem 18-5

Suppose an antenna has a purely resistive impedance of 100 Ω. It is connected to a ¼-wave section of 75-Ω coaxial cable. What is the impedance at the input end of the section?

Use the formula from above:

$$\begin{aligned} R_{in} &= Z_o^2/R_{out} \\ &= 75^2/100 \\ &= 5625/100 \\ &= 56\ \Omega \end{aligned}$$

Problem 18-6

Consider an antenna known to have a purely resistive impedance of 600 Ω. You want to match it to the output of a radio transmitter designed to work into a 50.0-Ω pure resistance. What is the characteristic impedance needed for a quarter-wave matching section?

Use this formula:

$$\begin{aligned} Z^2 &= R_{in}R_{out} \\ &= 600 \times 50 \\ &= 30{,}000 \end{aligned}$$

Therefore:

$$Z_o = (30{,}000)^{1/2}$$
$$= 173\ \Omega$$

It may be difficult to find a commercially manufactured transmission line that has this particular characteristic impedance. Prefabricated lines come in standard Z_o values, and a perfect match might not be obtainable. In that case, the closest obtainable Z_o should be used. In this case, it would probably be 150 Ω. If nothing is available anywhere near the characteristic impedance needed for a quarter-wave matching section, then a coil-type transformer can be used instead.

What about Reactance?

Things are simple when there is no reactance in an ac circuit using transformers. But often, especially in RF antenna systems, pure resistance doesn't occur naturally. It has to be obtained by using inductors and/or capacitors to cancel the reactance out. The presence of reactance in a load makes a perfect match impossible with an impedance-matching transformer alone.

Recall that inductive and capacitive reactances are opposite in effect, and that their magnitudes can vary. If a load presents a complex impedance $R + jX$, it is possible to cancel the reactance X by deliberately introducing an equal and opposite reactance $-X$. This can be, and often is, done by connecting an inductor or capacitor in series with a load that contains reactance as well as resistance. The result is a pure resistance with a value equal to $(R + jX) - jX$, or simply R.

When wireless communications is contemplated over a wide band of frequencies, adjustable impedance-matching and reactance-canceling networks can be placed between the transmitter and the antenna system. Such a device is called a *transmatch* or an *antenna tuner.* These devices not only match the resistive portions of the transmitter and load impedances, but they can tune out reactances in the load. Transmatches are popular among amateur radio operators, who use equipment capable of operation from less than 2 MHz up to the highest known radio frequencies.

Quiz

Refer to the text in this chapter if necessary. A good score is 18 or more correct. Answers are in the back of the book.

1. In a step-up transformer,
   (a) the primary impedance is greater than the secondary impedance.
   (b) the secondary winding is right on top of the primary.
   (c) the primary voltage is less than the secondary voltage.
   (d) All of the above are true.

2. The capacitance between the primary and the secondary windings of a transformer can be minimized by
   (a) placing the windings on opposite sides of a toroidal core.
   (b) winding the secondary right on top of the primary.
   (c) using the highest possible frequency.
   (d) using a center tap on the balanced winding.

3. A transformer steps a voltage down from 117 V to 6.00 V. What is its primary-to-secondary turns ratio?
   - (a) 1:380
   - (b) 380:1
   - (c) 1:19.5
   - (d) 19.5:1
4. A step-up transformer has a primary-to-secondary turns ratio of 1:5.00. If 117 V rms appears at the primary, what is the ac rms voltage across the secondary?
   - (a) 23.4 V rms
   - (b) 585 V rms
   - (c) 117 V rms
   - (d) 2.93 kV rms
5. A transformer has a secondary-to-primary turns ratio of 0.167. This transformer is
   - (a) a step-up unit.
   - (b) a step-down unit.
   - (c) neither a step-up unit nor a step-down unit.
   - (d) a reversible unit.
6. Which of the following statements is false, concerning air cores compared with ferromagnetic cores?
   - (a) Air concentrates the magnetic lines of flux.
   - (b) Air works at higher frequencies than ferromagnetics.
   - (c) Ferromagnetics are lossier than air.
   - (d) A ferromagnetic-core transformer needs fewer turns of wire than an equivalent air-core transformer.
7. Eddy currents cause
   - (a) an increase in efficiency.
   - (b) an increase in coupling between windings.
   - (c) an increase in core loss.
   - (d) an increase in usable frequency range.
8. Suppose a transformer has an ac voltage of 117 V rms across its primary, and 234 V rms appears across its secondary. If this transformer is reversed (that is, connected backward), assuming that this be done without damaging the windings, what will be the voltage at the output?
   - (a) 234 V rms
   - (b) 468 V rms
   - (c) 117 V rms
   - (d) 58.5 V rms
9. The shell method of transformer winding
   - (a) provides maximum coupling.
   - (b) minimizes capacitance between windings.

(c) withstands more voltage than other winding methods.
(d) has windings far apart but along a common axis.

10. Which of these core types is best if you need a winding inductance of 1.5 H?
(a) Air core
(b) Ferromagnetic solenoid core
(c) Ferromagnetic toroid core
(d) Ferromagnetic pot core

11. An advantage of a toroid core over a solenoid core is the fact that
(a) the toroid works at higher frequencies.
(b) the toroid confines the magnetic flux.
(c) the toroid can work for dc as well as for ac.
(d) it is easier to wind the turns on a toroid.

12. High voltage is used in long-distance power transmission because
(a) it is easier to regulate than low voltage.
(b) the I^2R losses are minimized.
(c) the electromagnetic fields are strong.
(d) small transformers can be used.

13. In a household circuit, 234-V rms electricity usually has
(a) one phase.
(b) two phases.
(c) three phases.
(d) four phases.

14. In a transformer, a center tap often exists in
(a) the primary winding.
(b) the secondary winding.
(c) an unbalanced winding.
(d) a balanced winding.

15. An autotransformer
(a) can be adjusted automatically.
(b) has a center-tapped secondary.
(c) consists of a single tapped winding.
(d) is useful only for impedance matching.

16. Suppose a transformer has a primary-to-secondary turns ratio of 2.00:1. The input impedance is 300 Ω, purely resistive. What is the output impedance?
(a) 75 Ω, purely resistive
(b) 150 Ω, purely resistive
(c) 600 Ω, purely resistive
(d) 1200 Ω, purely resistive

17. Suppose a purely resistive input impedance of 50 Ω must be matched to a purely resistive output impedance of 450 Ω. The primary-to-secondary turns ratio of the transformer must be which of the following?
    (a) 9.00
    (b) 3.00
    (c) 1/3.00
    (d) 1/9.00

18. Suppose a quarter-wave matching section has a characteristic impedance of 75.0 Ω. The input impedance is 50.0 Ω, purely resistive. What is the output impedance?
    (a) 150 Ω, purely resistive
    (b) 125 Ω, purely resistive
    (c) 100 Ω, purely resistive
    (d) 113 Ω, purely resistive

19. Suppose a purely resistive impedance of 75 Ω must be matched to a purely resistive impedance of 300 Ω. A quarter-wave section would need to have
    (a) $Z_o = 188\ \Omega$.
    (b) $Z_o = 150\ \Omega$.
    (c) $Z_o = 225\ \Omega$.
    (d) $Z_o = 375\ \Omega$.

20. If there is reactance in the load to which a transformer is connected, then
    (a) the transformer will be destroyed.
    (b) a perfect impedance match cannot be obtained.
    (c) a center tap must be used in the secondary.
    (d) the turns ratio must be changed to obtain an impedance match.

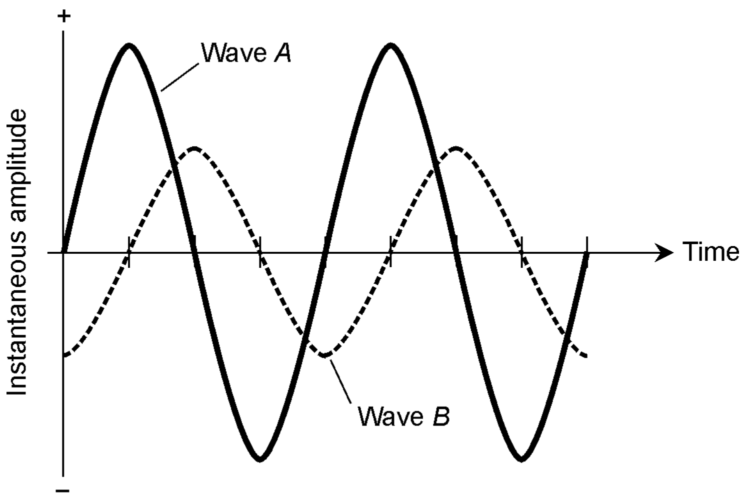
Test: Part 2

Do not refer to the text when taking this test. A good score is at least 37 correct. Answers are in the back of the book. It's best to have a friend check your score the first time, so you won't memorize the answers if you want to take the test again.

1. Consider a series circuit that has a resistance of 100 Ω and a capacitive reactance of −200 Ω. What is the complex impedance?
   (a) $-200 + j100$
   (b) $100 + j200$
   (c) $200 - j100$
   (d) $200 + j100$
   (e) $100 - j200$

2. Mutual inductance causes the net value of a set of coils to
   (a) cancel out, resulting in zero inductance.
   (b) be greater than what it would be with no mutual coupling.
   (c) be less than what it would be with no mutual coupling.
   (d) double.
   (e) vary, depending on the extent and phase of mutual coupling.

3. Refer to Fig. Test 2-1. Wave *A* is
   (a) leading wave *B* by 90°.
   (b) lagging wave *B* by 90°.
   (c) leading wave *B* by 180°.
   (d) lagging wave *B* by 135°.
   (e) lagging wave *B* by 45°.



Test 2-1 Illustration for Part 2 Test Question 3.

4. If a pure sine wave with no dc component has a positive peak value of +30.0 V pk, what is its rms voltage?
   (a) 21.2 V rms
   (b) 30.0 V rms
   (c) 42.4 V rms
   (d) 60.0 V rms
   (e) 90.0 V rms

5. Suppose four capacitors are connected in parallel. Their values are 100 pF each. What is the net capacitance?
   (a) 25 pF
   (b) 50 pF
   (c) 100 pF
   (d) 200 pF
   (e) 400 pF

6. Suppose an ac transformer has a primary-to-secondary turns ratio of 8.88/1. The input voltage is 234 V rms. What is the output voltage?
   (a) 2.08 kV rms
   (b) 18.5 kV rms
   (c) 2.97 V rms
   (d) 26.4 V rms
   (e) 20.8 V rms

7. In a series *RL* circuit, as the resistance becomes small compared with the reactance, the angle of lag approaches which of the following?
   (a) 0°
   (b) 45°

(c) 90°
(d) 180°
(e) 360°

8. Suppose an ac transmission line carries 3.50 A rms and 150 V rms. Imagine that the line is perfectly lossless, and that the load impedance is a pure resistance equal to the characteristic impedance of the line. What is the *true power* in this transmission line?
(a) 525 W
(b) 42.9 W
(c) 1.84 W
(d) Nonexistent, because true power is dissipated, not transmitted
(e) Variable, depending on standing-wave effects

9. In a parallel configuration, susceptances
(a) simply add up.
(b) add like capacitances in series.
(c) add like inductances in parallel.
(d) must be changed to reactances before you can work with them.
(e) cancel out.

10. Consider a sine wave that has a frequency of 200 kHz. How many degrees of phase change occur in a microsecond (a millionth of a second)?
(a) 180°
(b) 144°
(c) 120°
(d) 90°
(e) 72°

11. At a frequency of 2.55 MHz, what is the reactance of a 330-pF capacitor?
(a) −5.28 Ω
(b) −0.00528 Ω
(c) −189 Ω
(d) −18.9 kΩ
(e) −0.000189 Ω

12. Suppose a transformer has a step-up turns ratio of 1/3.16. The impedance of the load connected to the secondary is 499 Ω, purely resistive. What is the impedance at the primary?
(a) 50.0 Ω, purely resistive
(b) 158 Ω, purely resistive
(c) 1.58 kΩ, purely resistive
(d) 4.98 kΩ, purely resistive
(e) Impossible to calculate from the data given

13. If a complex impedance is represented by $34 - j23$, what is the absolute-value impedance?
   (a) 34 Ω
   (b) 11 Ω
   (c) −23 Ω
   (d) 41 Ω
   (e) 57 Ω

14. Suppose a coil has an inductance of 750 μH. What is the inductive reactance at 100 kHz?
   (a) 75.0 Ω
   (b) 75.0 kΩ
   (c) 471 Ω
   (d) 47.1 kΩ
   (e) 212 Ω

15. If two sine waves are 180° out of phase, it represents a difference of
   (a) ⅛ of a cycle.
   (b) ¼ of a cycle.
   (c) ½ of a cycle.
   (d) 1 full cycle.
   (e) 2 full cycles.

16. If *R* denotes resistance and *Z* denotes absolute-value impedance, then *R*/*Z* represents the
   (a) true power.
   (b) imaginary power.
   (c) apparent power.
   (d) absolute-value power.
   (e) power factor.

17. Suppose two components are connected in series. One component has a complex impedance of $30 + j50$, and the other component has a complex impedance of $50 - j30$. What is the impedance of the series combination?
   (a) $80 + j80$
   (b) $20 + j20$
   (c) $20 - j20$
   (d) $-20 + j20$
   (e) $80 + j20$

18. Suppose two inductors, having values of 140 μH and 1.50 mH, are connected in series. What is the net inductance?
   (a) 141.5 μH
   (b) 1.64 μH
   (c) 0.1415 mH

(d) 1.64 mH
(e) 0.164 mH

19. Which of the following types of capacitor is polarized?
(a) Mica
(b) Paper
(c) Electrolytic
(d) Air variable
(e) Ceramic

20. A coil with a toroidal, ferromagnetic core
(a) has less inductance than an air-core coil with the same number of turns.
(b) is essentially self-shielding.
(c) works well as a loopstick antenna.
(d) is ideal as a transmission-line transformer.
(e) cannot be used at frequencies below 10 MHz.

21. The efficiency of an electric generator
(a) depends on the mechanical driving power source.
(b) is equal to the electrical output power divided by the mechanical input power.
(c) depends on the nature of the electrical load.
(d) is equal to driving voltage divided by output voltage.
(e) is equal to driving current divided by output current.

22. Admittance is
(a) the reciprocal of reactance.
(b) the reciprocal of resistance.
(c) a measure of the opposition a circuit offers to ac.
(d) a measure of the ease with which a circuit passes ac.
(e) another expression for absolute-value impedance.

23. The absolute-value impedance Z of a parallel RLC circuit, where R is the resistance and X is the net reactance, is found according to which of the following formulas?
(a) $Z = R + X$
(b) $Z^2 = R^2 + X^2$
(c) $Z^2 = R^2X^2/(R^2 + X^2)$
(d) $Z = 1/(R^2 + X^2)$
(e) $Z = R^2X^2/(R + X)$

24. Complex numbers are used to represent impedance because
(a) reactance cannot store power.
(b) reactance isn't a real physical thing.

(c) they provide a way to represent what happens in resistance-reactance circuits.
(d) engineers like to work with sophisticated mathematics.
(e) Forget it! Complex numbers are never used to represent impedance.

25. Which of the following (within reason) has no effect on the value, in farads, of a capacitor?
(a) The mutual surface area of the plates
(b) The dielectric constant of the material between the plates
(c) The spacing between the plates
(d) The amount of overlap between plates
(e) The frequency

26. The 0° phase point in an ac sine wave is usually considered to be the point in time at which the instantaneous amplitude is
(a) zero and negative-going.
(b) at its negative peak.
(c) zero and positive-going.
(d) at its positive peak.
(e) any value; it doesn't matter.

27. The inductance of a coil can be adjusted in a practical way by
(a) varying the frequency of the signal applied to the coil.
(b) varying the number of turns using multiple taps.
(c) varying the current in the coil.
(d) varying the wavelength of the signal applied to the coil.
(e) varying the voltage across the coil.

28. Power factor is defined as the ratio of
(a) true power to VA power.
(b) true power to imaginary power.
(c) imaginary power to VA power.
(d) imaginary power to true power.
(e) VA power to true power.

29. Consider a situation in which you want to match a feed line with $Z_o = 50\ \Omega$ to an antenna with a purely resistive impedance of 200 Ω. A quarter-wave matching section should have which of the following?
(a) $Z_o = 150\ \Omega$
(b) $Z_o = 250\ \Omega$
(c) $Z_o = 125\ \Omega$
(d) $Z_o = 133\ \Omega$
(e) $Z_o = 100\ \Omega$

30. The vector $40 + j30$ in the *RX* plane represents
    (a) 40 Ω of resistance and 30 μH of inductance.
    (b) 40 μH of inductance and 30 Ω of resistance.
    (c) 40 Ω of resistance and 30 Ω of inductive reactance.
    (d) 40 Ω of inductive reactance and 30 Ω of resistance.
    (e) 40 μH of inductive reactance and 30 Ω of resistance.
31. In a series *RC* circuit where $R = 300\ \Omega$ and $X_C = -30\ \Omega$,
    (a) the current leads the voltage by a few degrees.
    (b) the current leads the voltage by almost 90°.
    (c) the voltage leads the current by a few degrees.
    (d) the voltage leads the current by almost 90°.
    (e) the voltage leads the current by 90°.
32. In a step-down transformer,
    (a) the primary voltage is greater than the secondary voltage.
    (b) the purely resistive impedance across the primary is less than the purely resistive impedance across the secondary.
    (c) the secondary voltage is greater than the primary voltage.
    (d) the output frequency is higher than the input frequency.
    (e) the output frequency is lower than the input frequency.
33. Suppose a capacitor of 470 pF is in parallel with an inductor of 4.44 μH. What is the resonant frequency?
    (a) 3.49 MHz
    (b) 3.49 kHz
    (c) 13.0 MHz
    (d) 13.0 GHz
    (e) It cannot be calculated from the data given.
34. A pure sine wave contains energy at
    (a) only one specific frequency.
    (b) a specific frequency and its even harmonics.
    (c) a specific frequency and its odd harmonics.
    (d) a specific frequency and all its harmonics.
    (e) a specific frequency and its second harmonic only.
35. Inductive susceptance is
    (a) the reciprocal of inductance.
    (b) negative imaginary.
    (c) equivalent to capacitive reactance.
    (d) the reciprocal of capacitive susceptance.
    (e) positive imaginary.

36. The rate of change (derivative) of a pure sine wave is another pure sine wave that has the same frequency as the original wave, and
    (a) is in phase with the original wave.
    (b) is 180° out of phase with the original wave.
    (c) leads the original wave by 45°.
    (d) lags the original wave by 90°.
    (e) leads the original wave by 90°.

37. True power is equal to
    (a) VA power plus imaginary power.
    (b) imaginary power minus VA power.
    (c) the vector difference between VA and reactive power.
    (d) VA power; the two are the same thing.
    (e) 0.707 times the VA power.

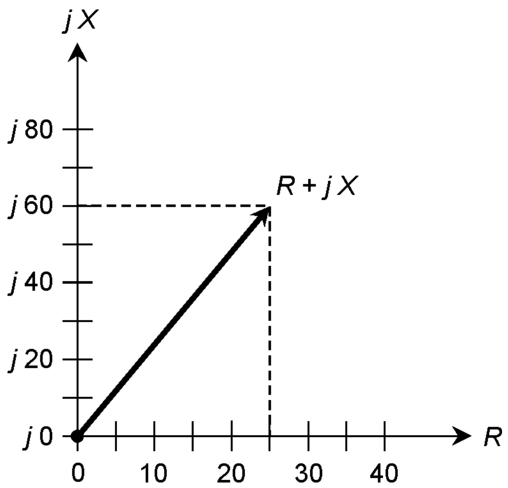
38. Consider a circuit in which three capacitors are connected in series. Their values are 47 μF, 68 μF, and 100 μF. The total capacitance of this combination is
    (a) 215 μF.
    (b) between 68 μF and 100 μF.
    (c) between 47 μF and 68 μF.
    (d) 22 μF.
    (e) not determinable from the data given.

39. The reactance of a section of transmission line depends on all of the following factors *except*
    (a) the velocity factor of the line.
    (b) the length of the section.
    (c) the current in the line.
    (d) the frequency of the signal in the line.
    (e) the wavelength of the signal in the line.

40. When analyzing a parallel *RLC* circuit to find the complex impedance, you should
    (a) add the resistance and reactance to get $R + jX$.
    (b) find the net conductance and susceptance, convert to resistance and reactance, and then add these to get $R + jX$.
    (c) find the net conductance and susceptance, and add these to get $R + jX$.
    (d) rearrange the components so they're connected in series, and find the complex impedance of that circuit.
    (e) subtract reactance from resistance to get $R - jX$.

41. The illustration in Fig. Test 2-2 shows a vector $R + jX$ representing
    (a) $X_C = 60\ \Omega$ and $R = 25\ \Omega$.
    (b) $X_L = 60\ \Omega$ and $R = 25\ \Omega$.
    (c) $X_L = 60\ \mu H$ and $R = 25\ \Omega$.

Test 2-2 Illustration for Part 2 Test Question 41.



(d) $C = 60\ \mu F$ and $R = 25\ \Omega$.
(e) $L = 60\ \mu H$ and $R = 25\ \Omega$.

42. Suppose two pure sine waves have no dc components, have the same frequency, and have the same peak-to-peak voltages, but they cancel each other out when combined. What is the phase difference between the waves?
(a) 45°
(b) 90°
(c) 180°
(d) 270°
(e) 360°

43. Suppose a series *RC* circuit has a resistance of 50 Ω and a capacitive reactance of −37 Ω. What is the phase angle?
(a) 37°
(b) 53°
(c) −37°
(d) −53°
(e) It cannot be calculated from the data given.

44. Suppose a 200-Ω resistor is in series with a coil and capacitor, such that $X_L = 200\ \Omega$ and $X_C = -100\ \Omega$. What is the complex impedance?
(a) $200 - j100$
(b) $200 - j200$
(c) $200 + j100$
(d) $200 + j200$
(e) Impossible to determine from the data given

45. The characteristic impedance of a transmission line
   (a) is negative imaginary.
   (b) is positive imaginary.
   (c) depends on the frequency.
   (d) depends on the construction of the line.
   (e) depends on the length of the line.

46. Suppose the period of a pure sine wave is 2×10^{-8}s. What is the frequency?
   (a) 2×10^{8} Hz
   (b) 20 MHz
   (c) 50 kHz
   (d) 50 MHz
   (e) 500 MHz

47. Suppose a series *RC* circuit has a resistance of 600 Ω and a capacitance of 220 pF. What is the phase angle?
   (a) −20°
   (b) 20°
   (c) −70°
   (d) 70°
   (e) Not determinable from the data given

48. A capacitor with a negative temperature coefficient
   (a) works less well as the temperature increases.
   (b) works better as the temperature increases.
   (c) heats up as its value is made larger.
   (d) cools down as its value is made larger.
   (e) exhibits increasing capacitance as the temperature drops.

49. Suppose three coils are connected in parallel. Each has an inductance of 300 μH. There is no mutual inductance. What is the net inductance?
   (a) 100 μH
   (b) 300 μH
   (c) 900 μH
   (d) 17.3 μH
   (e) 173 μH

50. Suppose a coil has 100 Ω of inductive reactance at 30.0 MHz. What is its inductance?
   (a) 0.531 μH
   (b) 18.8 mH
   (c) 531 μH
   (d) 18.8 μH
   (e) It can't be found from the data given.