Teach Yourself Electricity and Electronics

THIRD EDITION

STAN GIBILISCO

TAB ELECTRONICS

Teach Yourself Electricity and Electronics

This page intentionally left blank

Teach Yourself Electricity and Electronics

Third Edition

Stan Gibilisco

McGraw-Hill

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

To Tony, Tim, and Samuel
from Uncle Stan

This page intentionally left blank

Contents

Preface

This book is for people who want to learn basic electricity, electronics, and communications concepts without taking a formal course. It can also serve as a classroom text. This third edition contains new material covering acoustics, audio, high-fidelity, robotics, and artificial intelligence.

I recommend you start at the beginning of this book and go straight through. There are hundreds of quiz and test questions to fortify your knowledge and help you check your progress as you work your way along.

There is a short multiple-choice quiz at the end of every chapter. You may (and should) refer to the chapter texts when taking these quizzes. When you think you're ready, 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. The answers are listed in the back of the book. Stick with a chapter until you get most of the answers correct. Because you're allowed to look at the text during quizzes, the questions are written so that you really have to think before you write down an answer. Some are rather difficult, but there are no trick questions.

This book is divided into four major sections: Direct Current, Alternating Current, Basic Electronics, and Advanced Electronics and Related Technology. At the end of each section is a multiple-choice test. Take these tests when you're done with the respective sections and have taken all the chapter quizzes. Don't look back at the text when taking these tests. A satisfactory score is 37 answers correct. Again, answers are in the back of the book.

There is a final exam at the end of the book. The questions are practical, mostly nonmathematical, and somewhat easier than those in the quizzes. The final exam contains questions drawn from all the chapters. Take this exam when you have finished all four sections, all four section tests, and all of the chapter quizzes. A satisfactory score is at least 75 percent correct answers.

With the section tests and final exam, as with the quizzes, have a friend tell you your score without letting you know which questions you missed. That way, you will not subconsciously memorize the answers. You might want to take a test two or three times. When you have gotten a score that makes you happy, you can check to see where your knowledge is strong and where it can use some bolstering.

It is not necessary to have a mathematical or scientific background to use this do-it-yourself course. Junior-high-school algebra, geometry, and physical science will suffice. I've tried to gradually introduce standard symbols and notations so it will be evident what they mean as you go. By the time you get near the end of this book, assuming you've followed it all along, you should be familiar with most of the symbols used in schematic diagrams.

I recommend that you complete one chapter a week. An hour daily ought to be more than enough time for this. That way, in less than nine months, you'll complete the course. You can then use this book, with its comprehensive index, as a permanent reference.

Suggestions for future editions are welcome.

Stan Gibilisco

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 about *quantitative* things, the "what" versus the "how much." For now, you need only be concerned 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.

Even people of ancient times suspected that matter is made of invisible particles. They deduced this from observing things like water, rocks, and metals. These substances are much different from each other. But any given material—copper, for example—is the same wherever it is found. Even without doing any complicated experiments, early physicists felt that substances could only have these consistent behaviors if they were made of unique types, or arrangements, of particles. It took centuries before people knew just how this complicated business works. And even today, there are certain things that scientists don't really know. For example, is there a smallest possible material particle?

There were some scientists who refused to believe the atomic theory, even around the year of 1900. Today, practically everyone accepts the theory. It explains the behavior of matter better than any other scheme.

Eventually, scientists identified 92 different kinds of fundamental substances in nature, and called them *elements.* Later, a few more elements were artificially made.

Each 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 the atomic number

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.

Table 1-1 lists all the known elements in alphabetical order, with atomic numbers in one column, and atomic weights of the most common isotopes in another column. The standard abbreviations are also shown.

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).

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 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 neutrons, 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

Table 1-1. Atomic numbers and weights.

Element name	Abbreviation	Atomic number	Atomic weight*
Actinium	Ac	89	227
Aluminum	Al	13	27
Americium**	Am	95	243
Antimony	Sb	51	121
Argon	Ar	18	40
Arsenic	As	33	75
Astatine	At	85	210
Barium	Ba	56	138
Berkelium**	Bk	97	247
Beryllium	Be	4	9
Bismuth	Bi	83	209
Boron	B	5	11
Bromine	Br	35	79
Cadmium	Cd	48	114
Calcium	Ca	20	40
Californium**	Cf	98	251
Carbon	C	6	12
Cerium	Ce	58	140
Cesium	Cs	55	133
Chlorine	Cl	17	35
Chromium	Cr	24	52
Cobalt	Co	27	59
Copper	Cu	29	63
Curium**	Cm	96	247
Dysprosium	Dy	66	164
Einsteinium**	Es	99	254
Erbium	Er	68	166
Europium	Eu	63	153
Fermium	Fm	100	257
Fluorine	F	9	19
Francium	Fr	87	223
Gadolinium	Gd	64	158
Gallium	Ga	31	69
Germanium	Ge	32	74
Gold	Au	79	197
Hafnium	Hf	72	180
Helium	He	2	4
Holmium	Ho	67	165
Hydrogen	H	1	1
Indium	In	49	115
Iodine	I	53	127
Iridium	Ir	77	193
Iron	Fe	26	56
Krypton	Kr	36	84
Lanthanum	La	57	139
Lawrencium**	Lr or Lw	103	257

Table 1-1. *Continued*

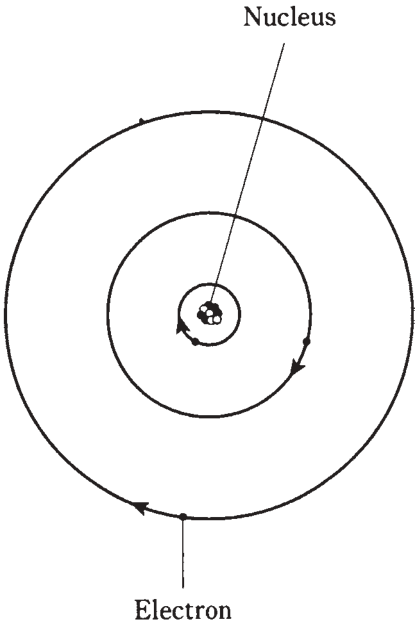
Element name	**Abbreviation**	**Atomic number**	**Atomic weight***
Lead	Pb	82	208
Lithium	Li	3	7
Lutetium	Lu	71	175
Magnesium	Mg	12	24
Manganese	Mn	25	55
Mendelevium**	Md	101	256
Mercury	Hg	80	202
Molybdenum	Mo	42	98
Neodymium	Nd	60	142
Neon	Ne	10	20
Neptunium**	Np	93	237
Nickel	Ni	28	58
Niobium	Nb	41	93
Nitrogen	N	7	14
Nobelium**	No	102	254
Osmium	Os	76	192
Oxygen	O	8	16
Palladium	Pd	46	108
Phosphorus	P	15	31
Platinum	Pt	78	195
Plutonium**	Pu	94	242
Polonium	Po	84	209
Potassium	K	19	39
Praseodymium	Pr	59	141
Promethium	Pm	61	145
Protactinium	Pa	91	231
Radium	Ra	88	226
Radon	Rn	86	222
Rhenium	Re	75	187
Rhodium	Rh	45	103
Rubidium	Rb	37	85
Ruthenium	Ru	44	102
Samarium	Sm	62	152
Scandium	Sc	21	45
Selenium	Se	34	80
Silicon	Si	14	28
Silver	Ag	47	107
Sodium	Na	11	23
Strontium	Sr	38	88
Sulfur	S	16	32
Tantalum	Ta	73	181
Technetium	Tc	43	99
Tellurium	Te	52	130
Terbium	Tb	65	159
Thallium	Tl	81	205
Thorium	Th	90	232
Thulium	Tm	69	169

Table 1-1. *Continued*

Element name	Abbreviation	Atomic number	Atomic weight*
Tin	Sn	50	120
Titanium	Ti	22	48
Tungsten	W	74	184
Unnilhexium**	Unh	106	—
Unnilpentium**	Unp	105	—
Unnilquadium**	Unq	104	—
Uranium	U	92	238
Vanadium	V	23	51
Xenon	Xe	54	132
Ytterbium	Yb	70	174
Yttrium	Y	39	89
Zinc	Zn	30	64
Zirconium	Zr	40	90

*Most common isotope. The sum of the number of protons and the number of neutrons in the nucleus. Most elements have other isotopes with different atomic weights.

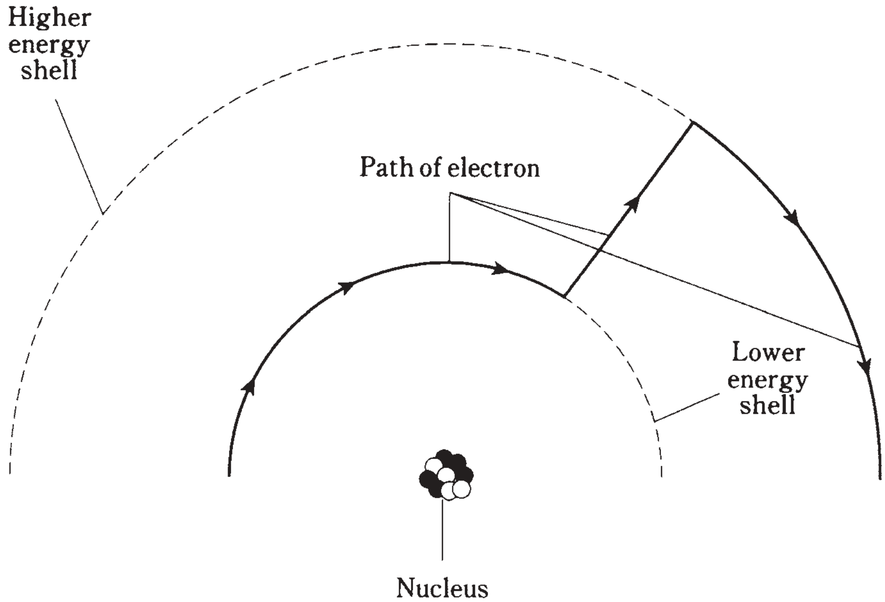
**These elements (atomic numbers 93 or larger) are not found in nature, but are human-made.



1-1 An early model of the atom, developed about the year 1900, rendered electrons like planets and the nucleus like the sun in a miniature solar system. Electric charge attraction kept the electrons from flying away.

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*.



1-2 Electrons move around the nucleus of an atom at defined levels corresponding to different energy states. This is a simplified drawing, depicting an electron gaining energy.

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.

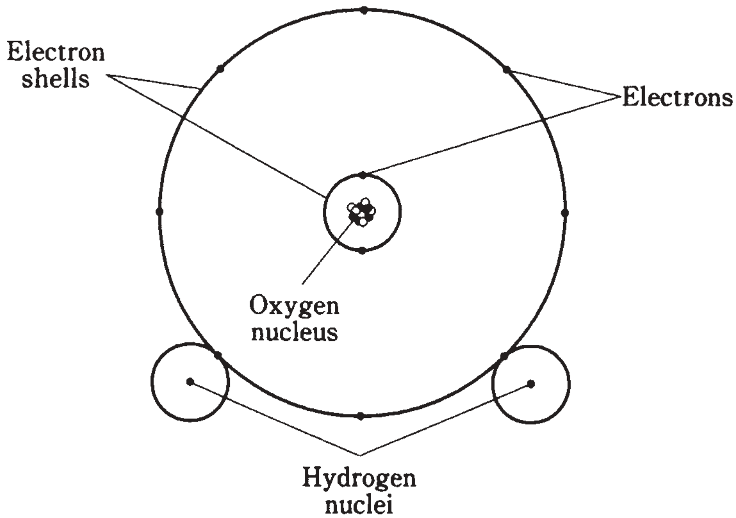
Ionized materials generally conduct electricity quite well, even if the substance is normally not a good conductor. Ionized air makes it possible for a lightning stroke to take place, for example. The ionization, caused by a powerful electric field, occurs along a jagged, narrow channel, as you have surely seen. 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.



1-3 Simplified diagram of a water molecule.

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.

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 are 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,* that has received much attention lately in environmental news. It is written O_3.

All matter, whether it is 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 are literally whizzing all over the place, bumping into each other and into solids and liquids adjacent to the gas (Fig. 1-4C).

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 are, in general, 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 is 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.

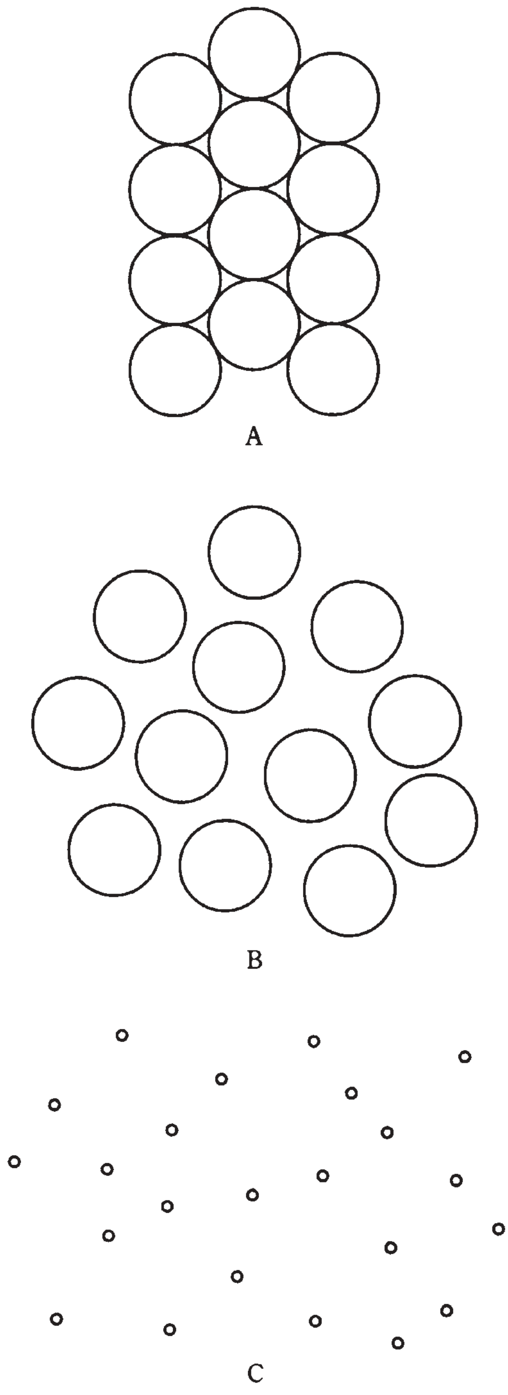
You might imagine a long line of people, each one constantly passing a ball to the neighbor on the right. If there are plenty of balls all along the line, and if everyone keeps passing balls along as they come, the result will be a steady stream of balls moving along the line. This represents a good conductor.

If the people become tired or lazy, and do not feel much like passing the balls along, the rate of flow will decrease. The conductor is no longer very good.

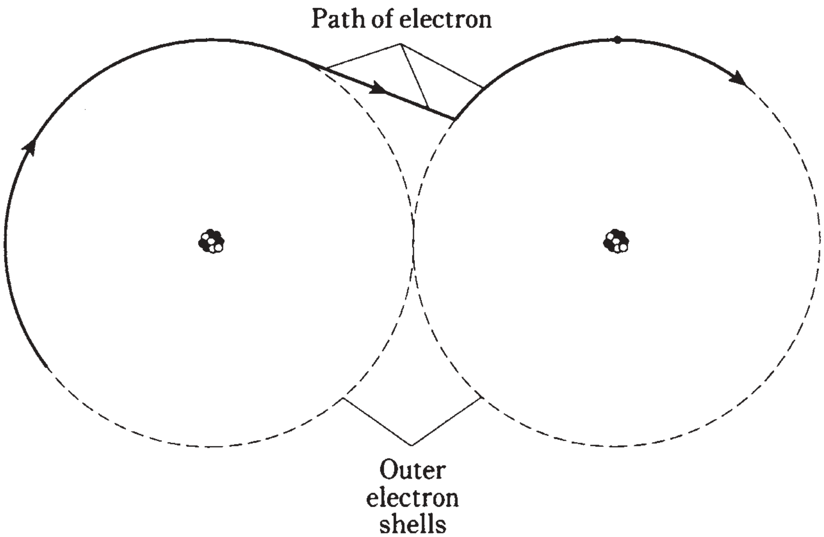
Insulators

If the people refuse to pass balls along the line in the previous example, the line represents an electrical *insulator.* Such substances prevent electrical currents from flowing, except possibly in very small amounts.

Most gases are good electrical insulators. Glass, dry wood, paper, and plastics are other examples. Pure water is a 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.



1-4 At A, simplified rendition of molecules in a solid; at B, in a liquid; at C, in a gas. The molecules don't shrink in the gas. They are shown smaller because of the much larger spaces between them.



1-5 In a conductor, electrons are passed from atom to atom.

Electrical insulators can be forced to carry current. Ionization can take place; when electrons are stripped away from their atoms, they have no choice but to move along. 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.

Resistors can be manufactured to have exact characteristics. Imagine telling each person in the line that they must pass a certain number of balls per minute. This is analogous to creating a resistor with a certain value of electrical *resistance.*

The better a resistor conducts, the lower its resistance; the worse it conducts, the higher the resistance.

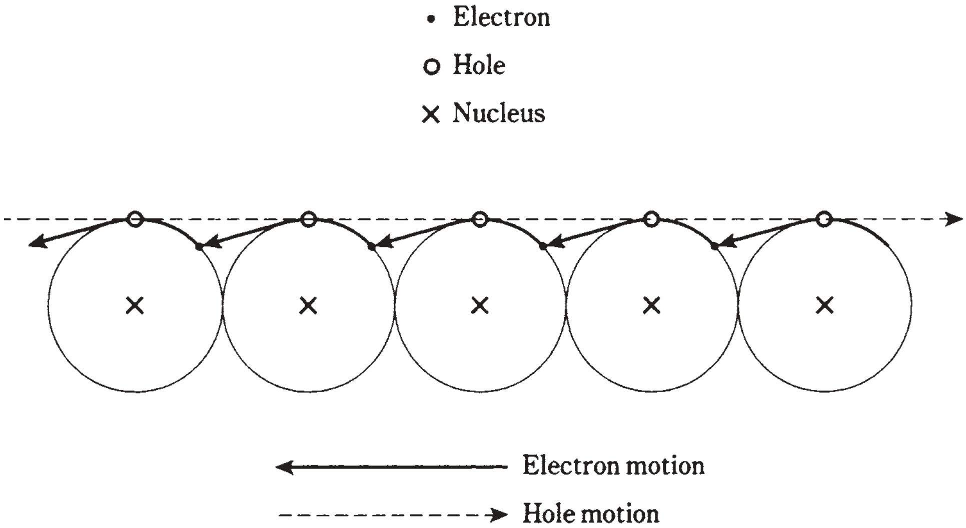
Electrical resistance is measured in units called *ohms.* The higher the value in ohms, the greater the resistance, and the more difficult it becomes for current to flow. For wires, the resistance is sometimes specified in terms of *ohms per foot* or *ohms per kilometer.* In an electrical system, it is usually desirable to have as low a resistance, or ohmic value, as possible. This is because resistance converts electrical energy into heat. Thick wires and high voltages reduce this resistance loss in long-distance electrical lines. This is why such gigantic towers, with dangerous voltages, are necessary in large utility systems.

Semiconductors

In a *semiconductor,* electrons flow, but not as well as they do in a conductor. You might imagine the people in the line being lazy and not too eager to pass the balls along. Some semiconductors carry electrons almost as well as good electrical conductors like copper or aluminum; others are almost as bad as insulating materials. The people might be just a little sluggish, or they might be almost asleep.

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

The semiconductors include certain substances, such as silicon, selenium, or gallium, that have been "doped" by the addition of impurities like indium or antimony. Perhaps you have 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. However, 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).



1-6 Holes move in the opposite direction from electrons in a semiconducting material.

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 semiconducting 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 in almost limitless variety. These substances are what make it possible for you to have a computer in a briefcase. That notebook computer, if it used vacuum tubes, would occupy a skyscraper, because it has billions of electronic components. It would also need its own power plant, and would cost thousands of dollars in electric bills every day. But the circuits are etched microscopically onto semiconducting wafers, greatly reducing the size and power requirements.

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 one second.

Usually, a great many charge carriers go past any given point in one 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 zeroes) charge carriers per second. Even the smallest mini-bulb carries *quadrillions* (numbers followed by 15 zeroes) of charge carriers every second. It is ridiculous to speak of a current in terms of charge carriers per second, so usually 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 one 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 one ampere of current.

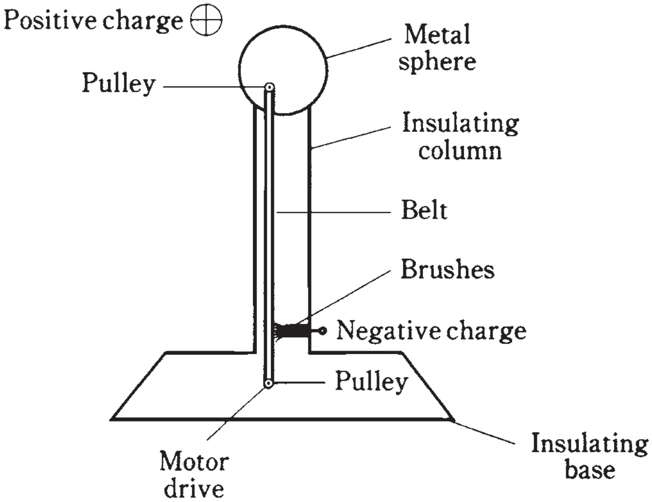
When a current flows through a resistance—and this is always the case because even the best conductors have resistance—heat is generated. Sometimes light and other forms of energy are emitted as well. A light bulb is deliberately designed so that the resistance causes visible light to be generated. Even the best incandescent lamp is inefficient, creating more heat than light energy. Fluorescent lamps are better. They produce more light for a given amount of current. Or, to put it another way, they need less current to give off a certain amount of light.

Electric current flows very fast through any conductor, resistor, or semiconductor. In fact, for most practical purposes you can consider the speed of current to be the same as the speed of light: 186,000 miles per second. Actually, it is a little less.

Static electricity

Charge carriers, particularly electrons, can build up, or become deficient, on things without flowing anywhere. You've probably experienced this when walking on a carpeted floor during the winter, or in a place where the humidity was very 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 that might well startle you. It is the current, during this discharge, that causes the sensation that might make you jump.

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, two inches, or even six 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 some high school physics labs, can cause a spark this large (Fig. 1-7). You have to be careful when using this device for physics experiments.



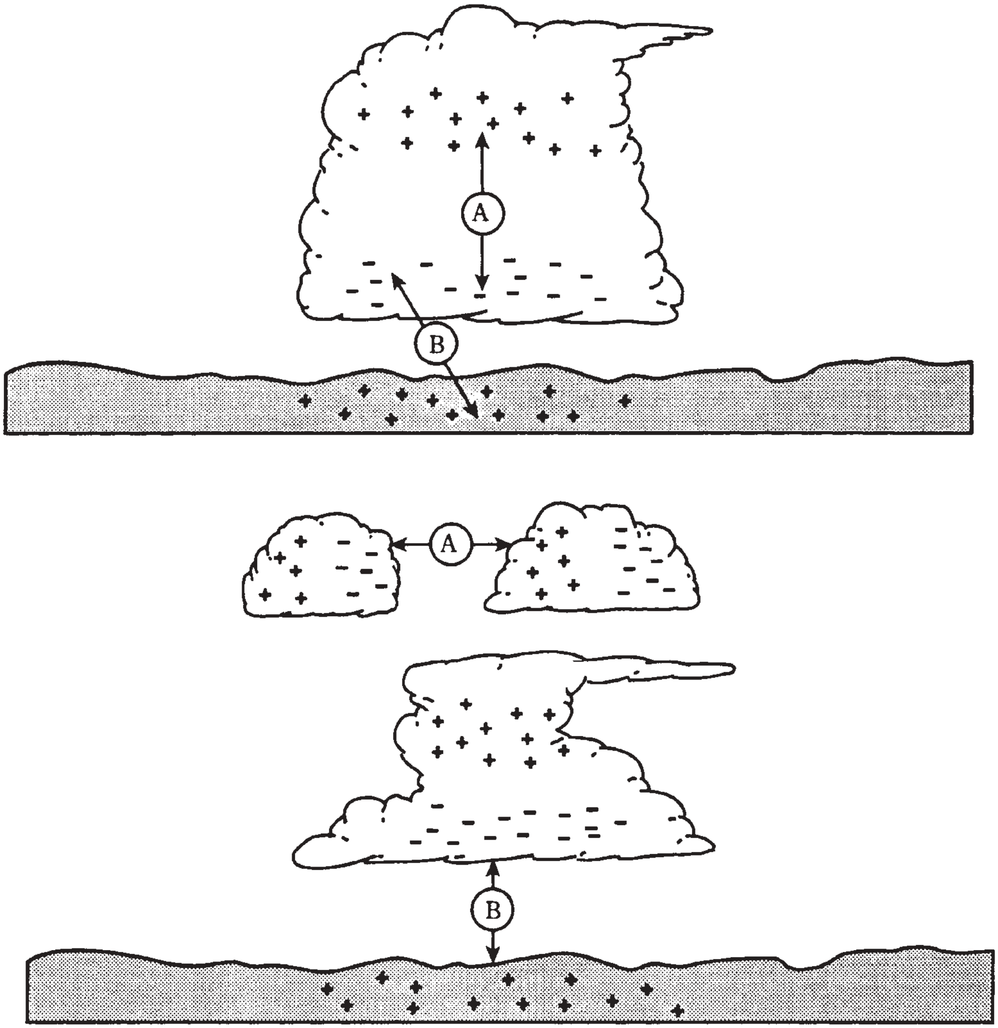
1-7 Simple diagram of a Van de Graaff generator for creating large static charges.

In the extreme, lightning occurs between clouds, and between clouds and ground in the earth's atmosphere. This spark is just a greatly magnified version of the little spark you get after shuffling around on a carpet. Until the spark 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 thunderstorm cloud like a shadow as the storm is blown along by the prevailing winds.

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 might be caused by a buildup of static electric charges, as in the case of a lightning stroke. When the charge builds up, with posi-



1-8 Cloud-to-cloud (A) and cloud-to-ground (B) charge buildup can both occur in a single thunderstorm.

tive polarity (shortage of electrons) in one place and negative polarity (excess of electrons) in another place, a powerful *electromotive force* exists. It is often abbreviated EMF. 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 volts (six volts in some older systems). 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, many millions of volts exist.

An EMF of one volt, across a resistance of one ohm, will cause a current of one 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 a little later in this book.

It is possible to have an EMF without having any current. This is the case just before a lightning bolt occurs, and before you touch that radiator after walking on the 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 very large EMF might not 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 that many coulombs of 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 so you don't get a severe shock.

Conversely, if there are plenty of coulombs available, a small voltage, such as 117 volts (or even less), can result in a lethal flow of current. This is why it is so 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 foolish enough to get caught in that kind of situation.

Nonelectrical energy

In electricity and electronics, there are many kinds of 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. And you are certainly acquainted with electric heaters, designed for the purpose of changing electricity into heat energy. This "heat" is actually a form of radiant energy called *infrared.* 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, and X rays. This is done by things like radio transmitters, sunlamps, and X-ray tubes.

Fast-moving protons, neutrons, electrons, and atomic nuclei are an important form of energy, especially in deep space where they are known as *cosmic radiation.* 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. Unfortunately, this form of energy, called *nuclear energy,* creates dangerous byproducts that are hard to dispose of.

When a conductor is moved in a magnetic field, electric current flows in that conductor. In this way, mechanical energy is converted into electricity. This is how a generator works. Generators can also work backwards. 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 flux 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 radio 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 all 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 awhile. 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. The resulting atomic weight is about:
   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. Might 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. Might have either a positive or negative charge.
6. A molecule:
   A. Might consist of just a single atom of an element.
   B. Must always contain two or more elements.
   C. Always has two or more atoms.
   D. Is always electrically charged.
7. In a compound:
   A. There can be just 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.
   B. By cooling.
   C. By ionizing.
   D. By oxidizing.
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:
    - 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 one ampere.
    - B. Flows through a 100-watt light bulb.
    - C. Is one 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 one volt is placed across a resistance of two ohms, then the current is:
    - A. Half an ampere.
    - B. One ampere.
    - C. Two amperes.
    - D. One ohm.
17. A backwards-working electric motor 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 generator.

18. In some batteries, chemical energy can 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 have to throw it away.
19. A changing 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. 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 SOME MORE ABOUT UNITS THAT QUANTIFY THE behavior of direct-current circuits. Many of these rules apply to utility alternating-current circuits also. Utility current is, in many respects, just like direct current because the frequency of alternation is low (60 complete cycles per second).

The volt

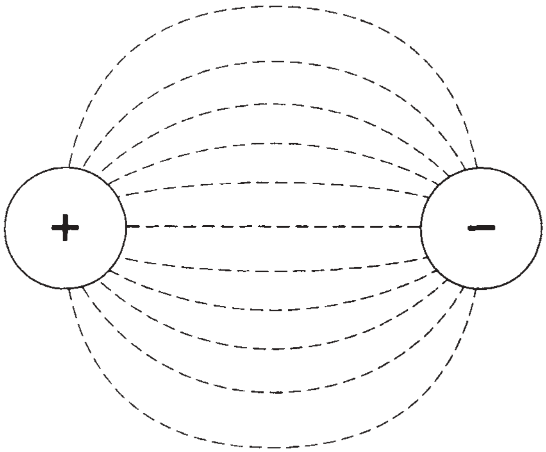
In chapter 1, you learned a little about the volt, the standard unit of electromotive force (EMF) or *potential difference*.

An accumulation of static electric charge, such as an excess or shortage of electrons, is always, associated with a voltage. There are other situations in which voltages exist. Voltage is generated at a power plant, and produced in an electrochemical reaction, and caused by light falling on a special 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 always 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 might not actually have a surplus of electrons with respect to neutral things. But there's always a difference in charge between the two poles. The negative pole always has more electrons than the positive pole.

The abbreviation for volt 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. And it is sometimes necessary to use units much larger than one volt. One kilovolt (kV) is equal to one thousand volts (1,000). One *megavolt* (MV) is equal to one million volts (1,000,000) or one thousand kilovolts.

In a dry cell, the EMF is usually between 1.2 and 1.7 V; in a car battery, it is most



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

often 12 V 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, kilovolts are used. The largest voltages on Earth occur between clouds, or between clouds and the ground, in thundershowers; this potential difference is on the order of tens of megavolts.

In every case, voltage, EMF, or potential difference represents the fact that charge carriers will flow between two points if a conductive path is provided. The number of charge carriers might be small even if the voltage is huge, or very large even if the voltage is tiny. Voltage represents the pressure or driving force that impels the charge carriers to move. In general, for a given number of charge carriers, higher voltages will produce a faster flow, and therefore a larger current. It's something like water pressure. The amount of water that will flow through a hose is proportional to the water pressure, all other things being equal.

Current flow

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

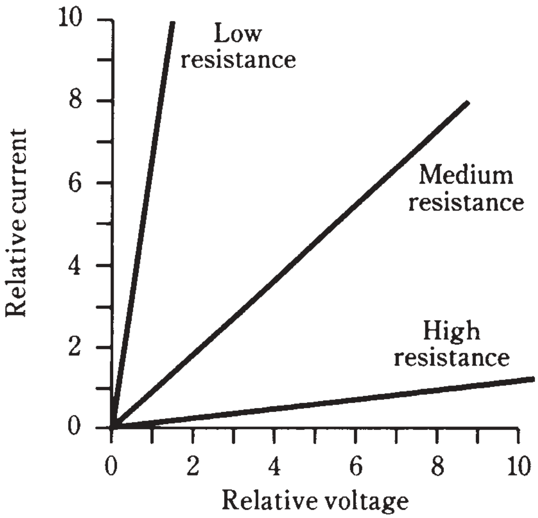
Sometimes the charge difference is equalized after a short while. This is the case, for example, when you touch a radiator after shuffling around on the carpet in your hard-soled shoes. It is also true in a lightning stroke. In these instances, the charge is equalized in a fraction of a second.

The charge might take longer to be used up. This will happen if you short-circuit a dry cell. Within a few minutes, or maybe up to an hour, the cell will "run 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 will essentially never equalize, 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.

You might have heard that "It's the current, not the voltage, that kills," concerning the danger in an electric circuit. 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.

Through an electric circuit with constant conductivity, the current is directly proportional to the applied voltage. That is, 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. But it assumes that the power supply can provide the necessary number of charge carriers. This rule holds only within reasonable limits.



2-2 Relative current versus relative voltage for different resistances.

When you are charged up with static electricity, there aren't very many charge carriers. A dry cell runs short of energy after awhile, and can no longer deliver as much current. All power supplies have their limitations in terms of the current they can provide. A power plant, or a power supply that works off of the utility mains, or a very large electrochemical battery, has a large capacity. You can then say that if you cut the resistance by a factor of 100, you'll get 100 times as much current. Or perhaps even 1000 or 10,000 times the current, if the resistance is cut to 0.001 or 0.0001 its former value.

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 per 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 = 0. 001 mA, a millionth of an ampere. And it is increasingly common to hear about *nanoamperes* (nA), where 1 nA = 0. 001 µA = 0.000000001 A (a billionth of an ampere). Rarely will you hear of *kiloamperes* (kA), where 1 kA = 1000 A.

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. An electric iron draws approximately 10 A; an entire household normally uses between 10 A 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 will flow 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, flow; this might 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 amounts of charge. 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 at all.

Resistance and the ohm

Resistance is a measure of the opposition that a circuit offers to the flow of electric current. You might 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 flow), and large-diameter wire has low resistance (not much opposition to electric currents). Of course, the type of metal makes a difference too. Iron wire has higher resistance for a given diameter than copper wire. Nichrome wire has still more resistance.

The standard unit of resistance is the ohm. This is sometimes abbreviated by the upper-case Greek letter omega, resembling an upside–down capital U (Ω). In this book, we'll just write it out as "ohm" or "ohms."

You'll sometimes hear about *kilohms* where 1 kilohm = 1,000 ohms, or about *megohms*, where 1 megohm = 1,000 kilohms = 1,000,000 ohms.

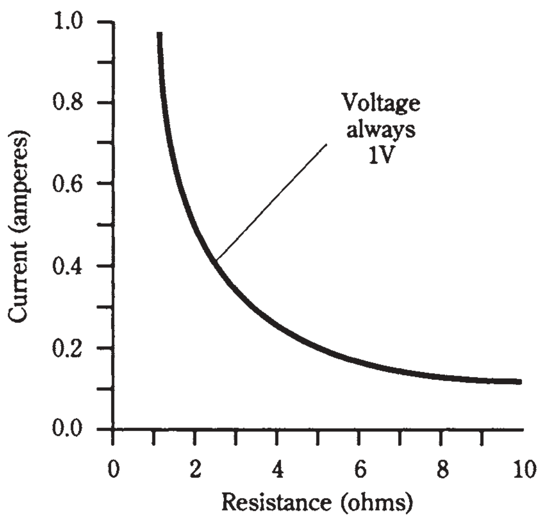
Electric wire is sometimes rated for *resistivity.* The standard unit for this purpose is the *ohm per foot* (ohm/ft) or the *ohm per meter* (ohm/m). You might also come across the unit *ohm per kilometer* (ohm/km). Table 2-1 shows the resistivity for various common sizes of wire.

When 1V is placed across 1 ohm of resistance, assuming that the power supply can

Table 2-1. Resistivity for copper wire, in terms of the size in American Wire Gauge (AWG).

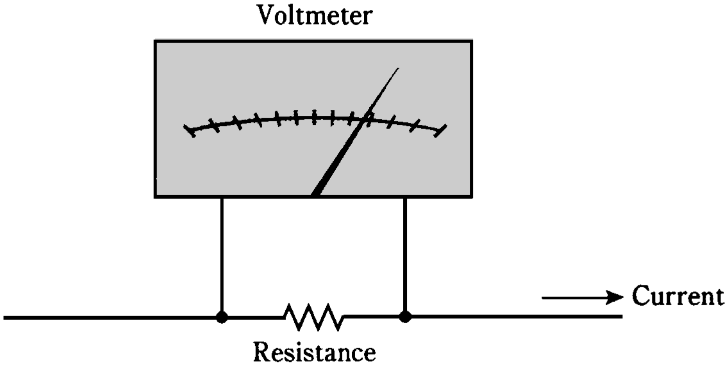
Wire size, AWG	Resistivity, ohms/km
2	0.52
4	0.83
6	1.3
8	2.7
10	3.3
12	5.3
14	8.4
16	13
18	21
20	34
22	54
24	86
26	140
28	220
30	350

deliver an unlimited number of charge carriers, there will be a current of 1A. If the resistance is doubled, the current is cut in half. If the resistance is cut in half, the current doubles. Therefore, the current flow, for a constant voltage, is *inversely proportional* to the resistance. Figure 2-3 is a graph that shows various currents, through various resistances, given a constant voltage of 1V across the whole resistance.



2-3 Current versus resistance through an electric device when the voltage is constant at 1 V.

Resistance has another property in an electric circuit. If there is a current flowing through a resistive material, there will always be a potential difference across the resistive object. This is shown in Fig. 2-4. The larger the current through the resistor, the greater the EMF across the resistor. In general, this EMF is directly proportional to the current through the resistor. This behavior of resistors is extremely useful in the design of electronic circuits, as you will learn later in this book.



2-4 Whenever a resistance carries a current, there is a voltage across it.

Electrical circuits always have some resistance. There is no such thing as a perfect conductor. When some metals are chilled to extremely low temperatures, 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.* In recent years, special metals have been found that behave this way even at fairly moderate temperatures. Researchers are trying to concoct substances that will superconduct even at room temperature. Superconductivity is an active field in physics right now.

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. These materials make good electric insulators, and good dielectrics for capacitors, devices that store electric charge.

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

Conductance and the siemens

The better a substance conducts, the less its resistance; the worse it conducts, the higher its resistance. Electricians and electrical engineers sometimes prefer to speak 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 ohm. 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 in ohms, you can get the conductance in siemens by taking the quotient of 1 over the resistance. Also, if you know the conductance in siemens, you can get the resistance in ohms by taking 1 over the conductance. The relation can be written as:

$$\text{siemens} = 1/\text{ohms, or}$$
$$\text{ohms} = 1/\text{siemens}$$

Smaller units of conductance are often necessary. A resistance of one kilohm is equal to one *millisiemens.* If the resistance is a megohm, the conductance is one *microsiemens.* You'll also hear about *kilosiemens or megasiemens,* representing resistances of 0.001 ohm and 0.000001 ohm (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 might sometimes have conductances in the megasiemens range.

As an example, suppose a component has a resistance of 50 ohms. Then its conductance, in siemens, is $1/50$, or 0.02 S. You might say that this is 20 mS. Or imagine a piece of wire with a conductance of 20 S. Its resistance is $1/20$, or 0.05, ohm. Not often will you hear the term "milliohm"; engineers do not, for some reason, speak of subohmic units very much. But you could say that this wire has a resistance of 50 milliohms, and you would be technically right.

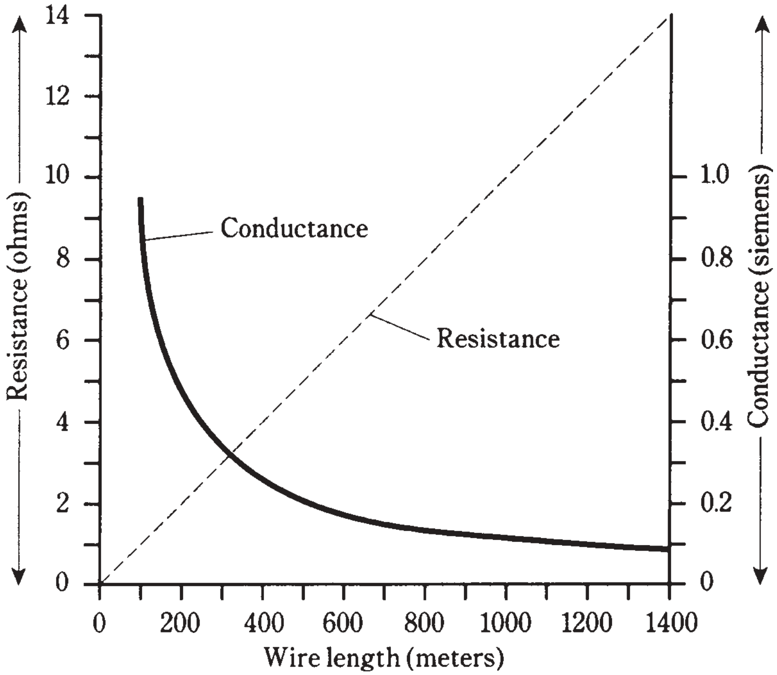
Conductivity is a little trickier. If wire has a resistivity of, say, 10 ohms per kilometer, you can't just say that it has a conductivity of $1/10$, or 0.1, siemens per kilometer. It is true that a kilometer of such wire will have a conductance of 0.1 S; but 2 km of the wire will have a resistance of 20 ohms (because there is twice as much wire), and this 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. Wrong! Conductance decreases, rather than increasing, with wire length.

When dealing with wire conductivity for various lengths of wire, it's best to convert to resistivity values, and then convert back to the final conductance when you're all done calculating. Then there won't be any problems with mathematical semantics.

Figure 2-5 illustrates the resistance and conductance values 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. This is inevitable. The heat can be measured in *watts,* abbreviated W, and represents electrical *power.* Power can be manifested in many other ways, such as in the form of mechanical motion, or radio waves, or visible light, or noise. In fact, there are dozens of different ways that power can be dissipated. 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.



2-5 Total resistances and conductances for a wire having 10 ohms of resistivity per kilometer.

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

This power might all be heat. Or it might exist in several forms, such as heat, light and infrared. This would be the state of affairs if the resistor were an incandescent light bulb, for example. If it were a motor, some of the power would exist in the form of mechanical work.

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

$$\text{(watts)} = EI = 3 \times 0.1 = 0.3 \text{ W}$$

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 = 0.855 A. Then

$$P \text{ (watts)} = 117 \times 0.855 = 100 \text{ W}$$

You will often hear about *milliwatts (mW), microwatts (μW), kilowatts (kW) and megawatts (MW)*. You should, by now, be able to tell from the prefixes what these units represent. But in case you haven't gotten the idea yet, you can refer to Table 2- 2. This table gives the most commonly used prefix multipliers in electricity and electronics, and the fractions that they represent. Thus, 1 mW = 0.001 W; 1 μW = 0.001 mW = 0.000001 W; 1 kW = 1,000 W; and 1 MW = 1,000 kW = 1,000, 000 W.

Table 2-2. Common prefix multipliers.

Prefix	**Fraction**
pico-	0.000000000001 (one-trillionth)
nano-	0.000000001 (one-billionth)
micro-	0.000001 (one-millionth)
milli-	0.001 (one-thousandth)
kilo-	1000
mega-	1,000,000
giga-	1,000,000,000 (one billion)
tera-	1,000,000,000,000 (one trillion)

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 power. It's easiest to remember that P = EI (watts equal volt-amperes), and derive the other equations from this by dividing through either by E (to get I) or by I (to get E).

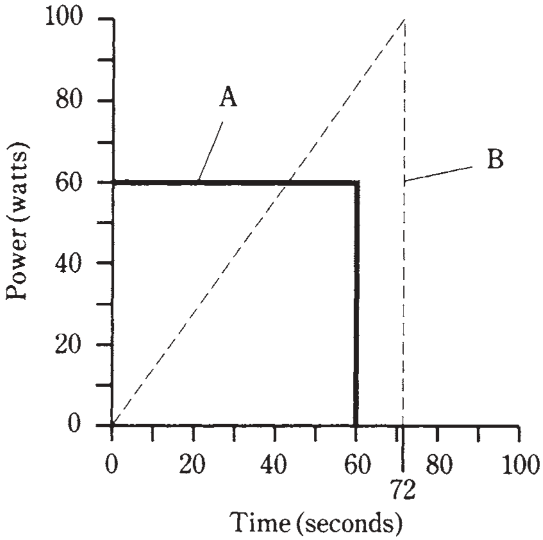
Energy and the watt hour

There is an important difference between energy and power. You've probably heard the two terms used interchangeably, as if they mean the same thing. But 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 *joules*. One joule is the equivalent of one watt of power, dissipated for one second of time. In electricity, you'll more often encounter the *watt hour* or the *kilowatt hour.* As their names imply, a watt hour, abbreviated Wh, is the equivalent of 1 W dissipated for an hour (1 h), and 1 kilowatt hour (kWh) is the equivalent of 1 kW of power dissipated for 1 h.

An energy of 1 Wh can be dissipated in an infinite number of different ways. A 60-watt bulb will burn 60 Wh in an hour, or 1 Wh per minute. A 100-W bulb would burn 1 Wh in $^{1}/_{100}$ hour, or 36 seconds. A 6-watt Christmas tree bulb would require 10 minutes ($^{1}/_{6}$ hour) to burn 1 Wh. And the rate of power dissipation need not be constant; it could be constantly changing.

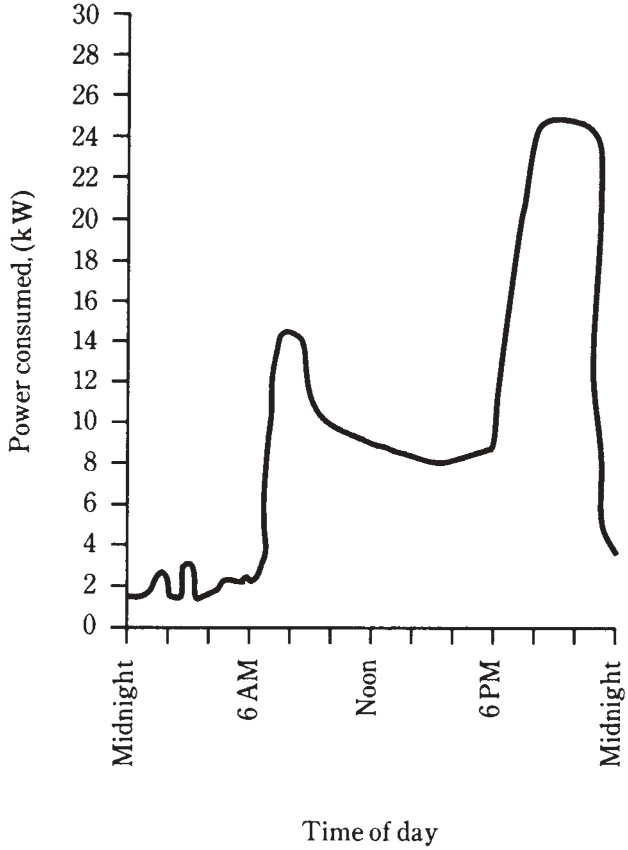
Figure 2-6 illustrates two hypothetical devices that burn up 1 Wh of energy. Device A uses its power at a constant rate of 60 watts, so it consumes 1 Wh in a minute. The power consumption rate of device B varies, starting at zero and ending up at quite a lot more than 60 W. How do you know that this second device really burns up 1 Wh of energy? You determine the area under the graph. This example has been chosen because figuring out this area is rather easy. Remember that the area of a triangle is equal to half the product of the base length and the height. This second device is on for 72 seconds, or 1.2 minute; this is $1.2/60 = 0.02$ hour. Then the area under the "curve" is $\frac{1}{2} \times 100 \times 0.02 = 1$ Wh.



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

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. That means a wrong answer!

Sometimes, the curves in graphs like these are complicated. In fact, they usually are. Consider the graph of power consumption in your home, versus time, for a whole day. It might look something like the curve in Fig. 2-7. Finding the area under this curve is no easy task, if you have only this graph to go by. But there is another way to determine the total energy burned by your household in a day, or in a week, or most often, in a month. That is by means of the electric meter. It measures electrical energy in kilowatt hours. Every month, without fail, the power company sends its representative to read that meter. This person takes down the number of kilowatt hours displayed, subtracts the number from 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 do sophisticated integral calculus to find the areas under irregular curves such as the graph of Fig. 2-7.



2-7 Hypothetical graph showing the power consumed by a typical household, as a function of the time of day.

Other energy units

As said before, physicists prefer to use the joule, or watt second, as their energy unit. This is the standard unit for scientific purposes.

Another unit is the *erg,* equivalent to one ten-millionth (0.0000001) of a joule. This is said to be roughly the amount of energy needed by a mosquito to take off after it has bitten you (not including the energy needed for the bite itself). The erg is used in lab experiments involving small amounts of expended energy.

You have probably heard of the *British thermal unit (Btu),* equivalent to 1055 joules. This is the energy unit most often used to indicate the cooling or heating capacity of air-conditioning equipment. To cool your room from 85 to 78 degrees 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, in Btu, that best suits your needs. It doesn't make any sense to get one that is way too big; you'll be wasting your money. But you want to be sure that it's big enough—or you'll also waste money because of inefficiency and possibly also because of frequent repair calls.

Physicists also use, in addition to the joule, a unit of energy called the *electron volt* (eV). This is an extremely 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. Atom smashers are rated by millions of electron volts (MeV) or billions of electron volts (GeV) of energy capacity. In the future you might even hear of a huge linear accelerator, built on some vast prairie, and capable of delivering trillions of electron volts (TeV).

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. You don't really need to worry about the exponential notation, called *scientific notation,* here. In electricity and electronics, you need to be concerned only with the watt hour and the kilowatt hour for most purposes, and the conversions hardly ever involve numbers so huge or so miniscule that you'll need scientific notation.

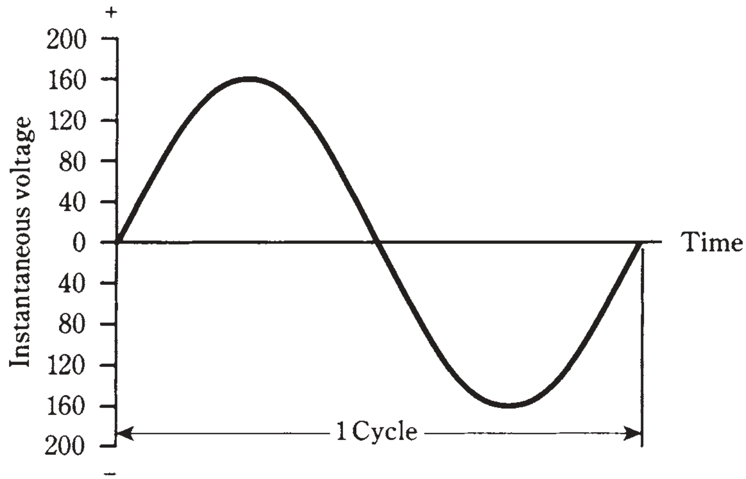
Table 2-3. Energy units.

Unit	To convert to joules multiply by	Conversely, multiply by
Btu	1055	0.000948 or 9.48×10^{-4}
eV	1.6×10^{-19}	6.2×10^{18}
erg	0.0000001 or 10^{-7}	10,000,000 or 10^{7}
ft-lb	1.356	0.738
Wh	3600	0.000278 or 2.78×10^{-4}
kWh	3,600,000 or 3.6×10^{6}	0.000000278 or 2.78×10^{-7}

ac Waves and the hertz

This chapter, and this whole first section, is concerned with *direct current* (dc), that is, 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). In some countries, the direction reverses every $^1/_{100}$ second, and the cycle is completed every $^1/_{50}$ second.

Figure 2-8 shows the characteristic wave of alternating current, 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 waveshape 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. The peak voltage is about 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 word "Hertz" literally translates to " cycles per second." In the U.S., this is the standard frequency for ac. In some places it is 50 Hz. (Some remote places even use dc, but they are definitely the exception, not the rule.)

In radio practice, higher frequencies are common, and you'll hear about *kilohertz (kHz), megahertz (MHz) and gigahertz (GHz).* You should know right away the size of these units, but in case you're still not sure about the way the prefixes work, the relationships are as follows:

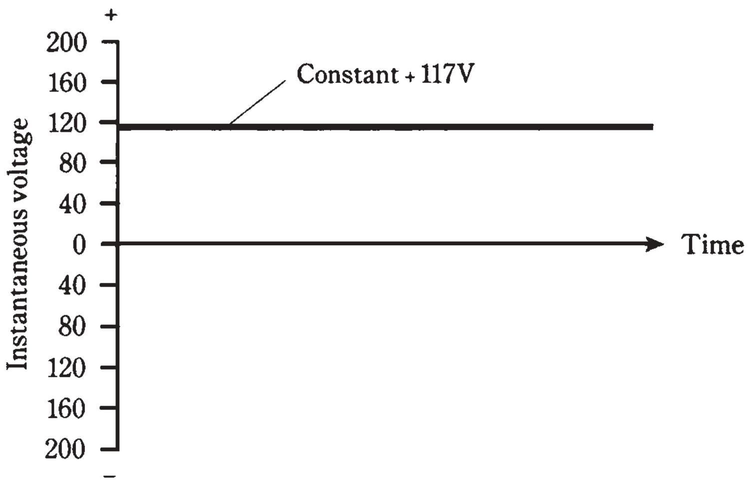
$$\begin{aligned} 1 \text{ kHz} &= 1000 \text{ Hz} \\ 1 \text{ MHz} &= 1000 \text{ kHz} = 1{,}000{,}000 \text{ Hz} \\ 1 \text{ GHz} &= 1000 \text{ MHz} = 1{,}000{,}000 \text{ kHz} \\ &= 1{,}000{,}000{,}000 \text{ Hz} \end{aligned}$$

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 fluctuating direct current

Batteries and other sources of direct current (dc) produce a constant voltage. This can be represented by a straight, horizontal line on a graph of voltage versus time (Fig. 2-9).

The peak and effective values are the same for pure dc. But sometimes the value of dc voltage fluctuates rapidly with time, in a manner similar to the changes in an ac wave. This might happen if the waveform in Fig. 2-8 is *rectified.*



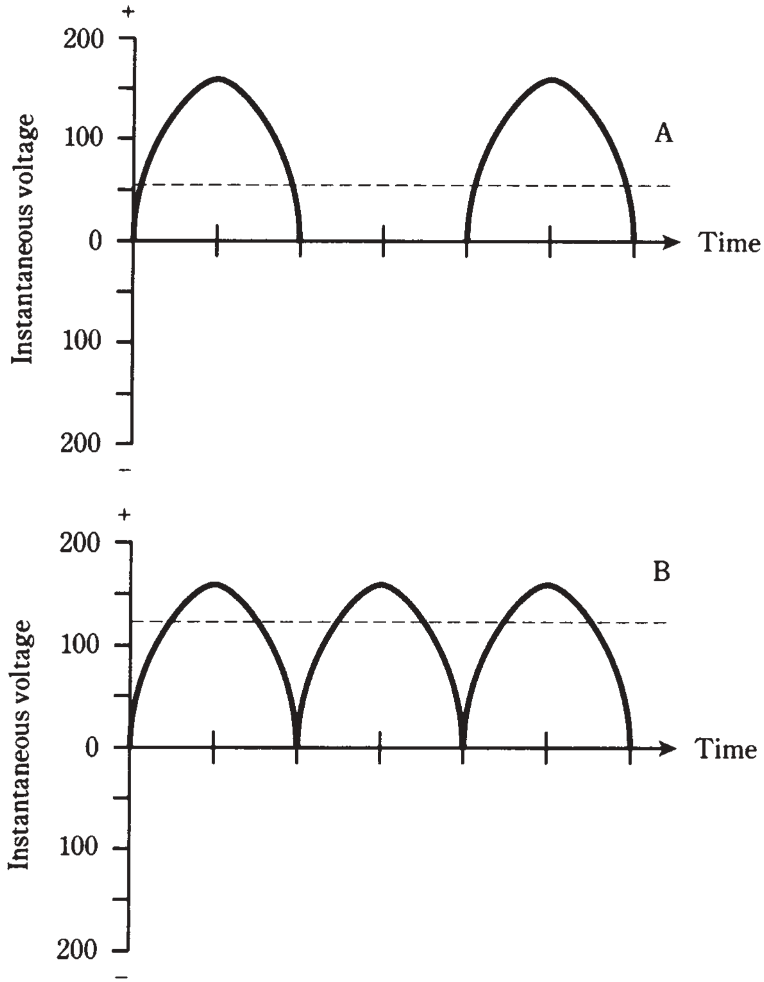
2-9 A representation of pure dc.

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 that the *polarity* is always the same, either positive or negative.

Figure 2-10 illustrates two different waveforms of fluctuating, or pulsating, dc. In the waveform at A, the negative (bottom) part has simply been chopped off. At B, the negative portion of the wave has been turned around and made positive, a mirror image of its former self. The situation at A is known as *half-wave rectification,* because it makes use of only half the wave. At B, the wave has been subjected to *full-wave rectification,* because 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 has been used. In the figure, effective voltage is shown as a dotted line, and the *instantaneous voltage is* shown as a solid line. Notice that the instantaneous voltage changes all the time, from instant to instant. This is how it gets this name! The peak voltage is the maximum instantaneous voltage. Instantaneous voltage is never, ever any greater than the peak.

In Fig. 2-10B, the effective value is 0.707 times the peak value, 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 just 0.354 times the peak value.



2-10 At A, half-wave rectification of ac. At B, full-wave rectification. Effective values are shown by dotted lines.

Using household ac as an example, the peak value is generally about 165 V; the effective value is 117 V. If full-wave rectification is used (Fig.2-10B), the effective value is still 117 V. If half-wave rectification is used, as in Fig. 2-10A, the effective value is about 58.5 V.

Safety considerations in electrical work

For your purposes here, one rule applies concerning safety around electrical apparatus. If you are in the slightest doubt about whether or not something is safe, *leave it to a professional electrician.*

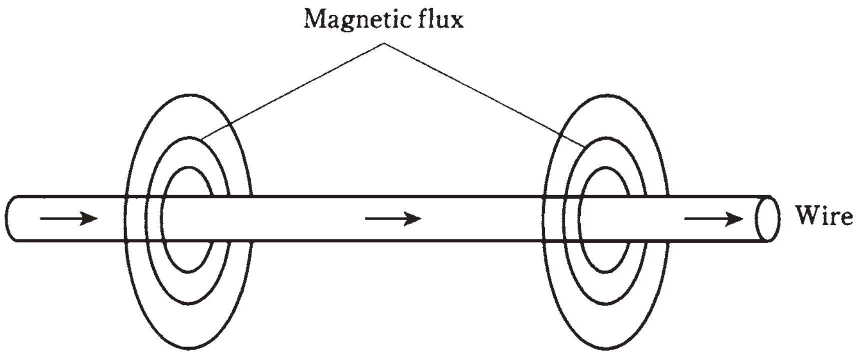
Household voltage, normally about 117 V but sometimes twice that, or about 234 V, is more than sufficient to kill you if it appears across your chest cavity. Certain devices, such as automotive spark coils, can produce lethal currents even from the low voltage (12 V to 14 V) in a car battery.

Consult the American Red Cross or your electrician concerning what kinds of circuits, procedures and devices are safe, and which kinds 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, the magnetic *lines of flux* surround the wire in circles, with the wire at the center (Fig. 2-11). Actually, these aren't really lines or circles; this is just a convenient way to represent the magnetic field. You might sometimes hear of 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.

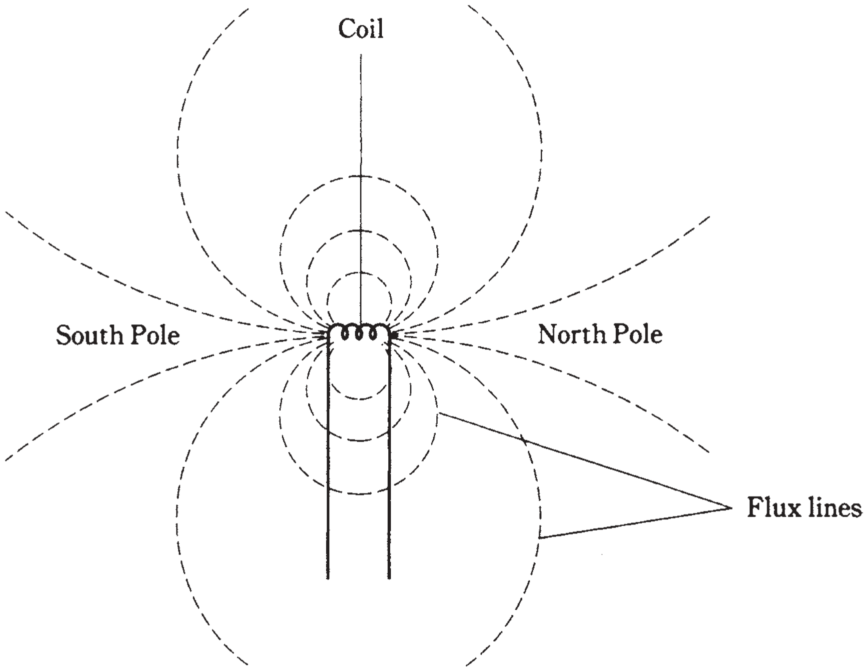


2-11 Magnetic flux lines around a straight, current-carrying wire. The arrows indicate current flow.

Magnetic fields can be produced when the atoms of certain materials align themselves. Iron is the most common metal that has this property. The iron in the core of the earth has 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



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

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. But this is just a semantical thing, about which theoretical physicists might speak. It doesn't need to concern you for ordinary electrical and electronics applications.

Magnetic units

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

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*. One *gauss* is equal to 0.0001 weber, or one maxwell per square centimeter.

In general, the greater the electric current through a wire, the greater the flux density near the wire. A coiled wire will produce a greater flux density 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, will double 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, again, 10 At. (Remember that 100 mA = 0.1 A.)

Another unit of magnetomotive force is the *gilbert.* This unit is equal to 0.796 At.

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 one volt:
   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 power plant.
   D. A clock radio.

5. A piece of wire has a conductance of 20 siemens. 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 ohms. Its conductance is:

   A. 3.33 millisiemens.

   B. 33.3 millisiemens.

   C. 333 microsiemens.

   D. 0.333 siemens.

7. A mile of wire has a conductance of 0.6 siemens. Then three miles of the same wire has a conductance of:

   A. 1.8 siemens.

   B. 1.8 Ω.

   C. 0.2 siemens.

   D. Not enough information has been given to answer this.

8. A 2-kW generator will deliver approximately how much current, reliably, 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. This represents approximately how many kilowatts?

   A. 1.76.

   B. 1760.

   C. 7.8.

   D. 0.0078.

10. You are told that a certain air conditioner is rated at 500 Btu. What is this in kWh?

    A. 147.

    B. 14.7.

    C. 1.47.

    D. 0.147.

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 U.S. 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 chopped off.
C. The whole wave is inverted.
D. The effective value is half the peak value.

14. In the output of a half-wave rectifier:
A. Half of the wave is inverted.
B. The effective value is less than that of the original ac wave.
C. The effective value is the same as that of the original ac wave.
D. The effective value is more than that of the original ac wave.

15. In the output of a full-wave rectifier:
A. The whole wave is inverted.
B. The effective value is less than that of the original ac wave.
C. The effective value is the same as that of the original ac wave.
D. The effective value is more than that of the original ac 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 these 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 will be:

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. By using a tension spring against which the electric or magnetic force can pull or push, a movable needle can be constructed. The needle can then be placed in front of a calibrated scale, allowing a direct reading of the quantity to be measured. These 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.

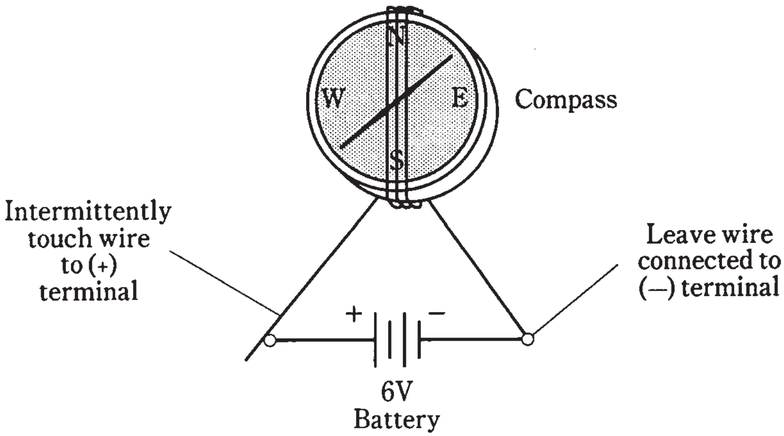
Some meters work by means of small motors whose speed depends on the measured quantity. The rotation rate, or the number of rotations in a given time, can be measured or counted. These are forms of *rate meters*.

Still other kinds of meters actually count electronic pulses, sometimes in thousands, millions or billions. These are *electronic counters*. There are also various other metering methods.

Electromagnetic deflection

Early experimenters with electricity and magnetism noticed that an electric current produces a magnetic field. This discovery was probably an accident, but it was an accident that, given the curiosity of the scientist, was bound to happen. 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 error depends on how close the compass is brought to the wire, and also on how much current the wire is carrying.

Scientific experimenters are like children. They like to play around with things. Most likely, when this effect was first observed, the scientist 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 would indicate a tiny electric current (Fig. 3-1). This effect is known as *galvanism,* and the meter so devised was called a *galvanometer.*

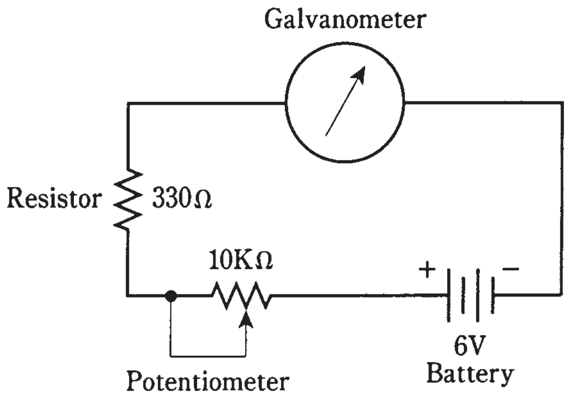


3-1 A simple galvanometer. The compass must lie flat.

Once this device was made, the scientist saw that the extent of the needle displacement increased with increasing current. Aha—a device for measuring current! Then, the only challenge was to calibrate the galvanometer somehow, and to set up some kind of standard so that a universal meter could be engineered.

You can easily make your own galvanometer. Just buy a cheap compass, about two feet of insulated bell wire, and a six-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 right along the wire turns while the wire is disconnected from the battery. Connect one end of the wire to the minus (–) terminal of the battery. Touch the other end to the plus (+) terminal, intermittently, 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 Radio Shack, 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 at ¼ watt, and the potentiometer should have a value of 10 KΩ (10,000 Ω) maximum. 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 degree scale around the perimeter of the compass.

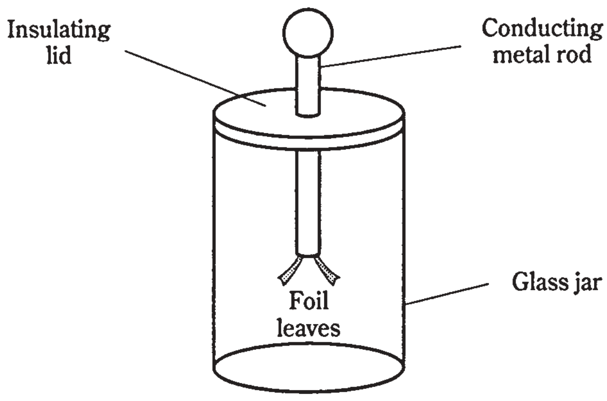


3-2 Circuit for demonstrating how a galvanometer indicates relative current.

Electrostatic deflection

Electric fields produce forces, just as do magnetic fields. You have probably noticed this when your hair feels like it's standing on end in very dry or cold weather. You've probably heard that people's hair really does stand straight out just before a lightning bolt hits nearby; this is no myth. Maybe you performed experiments in science classes to observe this effect.

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 will not 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.



3-3 A simple electroscope.

The extent to which the leaves stand apart depends on the amount of electric charge. It is somewhat difficult to actually 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 device has the ability to measure alternating electric charges as well as steady charges. This gives electrostatic meters an advantage over electromagnetic meters (galvanometers). If you connect ac to the coil of the galvanometer device in Fig. 3-1 or Fig. 3-2, the compass needle might vibrate, but will not give a clear deflection. This is because current in one direction pulls the meter needle one way, and current in the other direction will deflect the needle the opposite way. But if an alternating electric field is connected to an electrostatic meter, the plates will repel whether the charge is positive or negative. The deflection will be steady, therefore, with ac as well as with dc.

Most electroscopes aren't sensitive enough to show much deflection with ordinary 117-V utility voltage. Don't try connecting 117 V to an electroscope anyway; it might not deflect the foil leaves, but it can certainly present a danger to your body if you bring it out to points where you can readily 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 amount at first, needed to put a charge on the plates. Sometimes, an engineer or experimenter doesn't want the measuring device to draw current, because this affects the behavior of the circuit under test. Galvanometers, by contrast, always need at least a little bit of current in order to operate. You can observe this effect by charging up a laboratory electroscope, say with a glass rod that has been rubbed against a cloth. When the rod is pulled away from the electroscope, the foil leaves will remain standing apart. The charge just sits there. If the electroscope drew any current, the leaves would immediately 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 just 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.

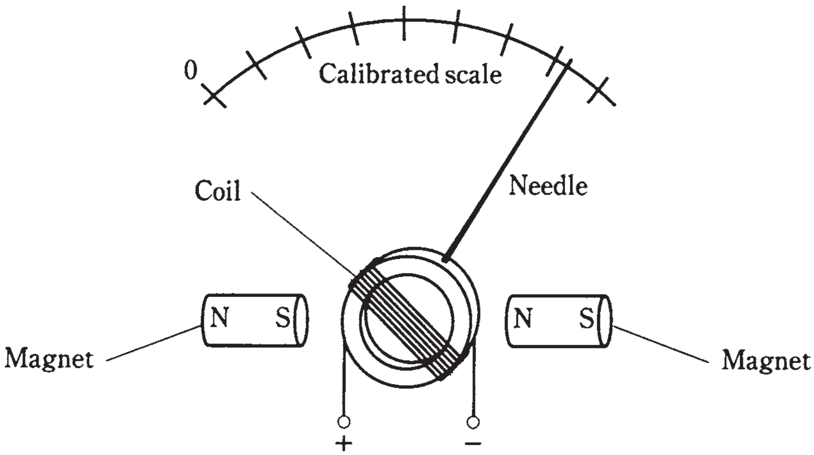
A variation of the hot-wire principle can be used by placing two different metals into contact with each other. If the right metals are chosen, the junction will heat up when a current flows through it. This is called the *thermocouple principle.* As with the hot-wire meter, a thermometer can be used to measure the extent of the heating.

But there is also another effect. A thermocouple, when it gets warm, generates a direct current. This current can be measured by a more conventional, dc type meter. This method is useful when it is necessary to have a faster meter response time. The hot-wire and thermocouple effects are used occasionally to measure current at radio frequencies, in the range of hundreds of kilohertz up to tens of gigahertz.

Ammeters

Getting back to electromagnetic deflection, and the workings of the galvanometer, you might have thought by now that a magnetic compass doesn't make a very convenient type of 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 turned in just the right way, so as to be aligned with the north geomagnetic pole. That would not only be a great bother, but it would be ridiculous. Imagine a bunch of scientists running around, turning radios and other apparatus so the meters are all lying flat and are all lined up with the earth's magnetic field! In the early days of electricity and electronics, when the phenomena were confined to scientific labs, this was indeed pretty much how things were.

Then someone thought that the magnetic field could be provided by a permanent magnet right inside the meter, instead of by the earth. This would supply a stronger magnetic force, and would therefore make it possible to detect much weaker currents. It would let the meter be turned in any direction and the operation would not be affected. The coil could be attached right to the meter pointer, and suspended by means of a spring in the field of the magnet. This kind of meter, called a *D'Arsonval movement,* is still extensively used today. The assembly is shown in Fig. 3-4. This is the basic principle of the *ammeter.*



3-4 The D'Arsonval meter movement. The spring bearing is not shown.

A variation of this is the attachment of the meter needle to a permanent magnet, and the winding of 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 meter movement is also sometimes called a D'Arsonval movement. This method will work, but the inertial mass of the permanent magnet causes a slower needle response. This kind 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 current level before finally coming to rest at the correct reading.

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 meter, and of its cousins, 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, and the smaller the amount of current that is needed to cause a certain amount of needle movement.

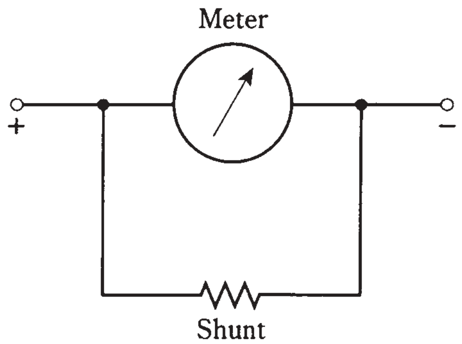
The most sensitive ammeters can detect currents of just a microampere or two. The amount of current for *full scale deflection* (the needle goes all the way up without banging against the stop pin) can be as little as about 50 uA in commonly available meters. Thus you might see a *microammeter,* or a *milliammeter,* quite often in electronic work. Meters that measure large currents are not a problem to make; it's easy to make an insensitive device.

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, since this 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 take 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 factor of 10, or 100, or even 1000. This resistor must be capable of carrying the current without burning up. It might have to take 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 frequently used when it is necessary to measure very large currents, such as hundreds of amperes. They allow microammeters or milliammeters to be used in a versatile *multimeter,* with many current ranges.

Voltmeters

Current is a flow of charge carriers. Voltage, or electromotive force (EMF), or potential difference, is the "pressure" that makes a current possible. Given a circuit whose resistance is constant, the current that will flow in the circuit is directly proportional to the



3-5 A resistor can be connected across a meter to reduce the sensitivity.

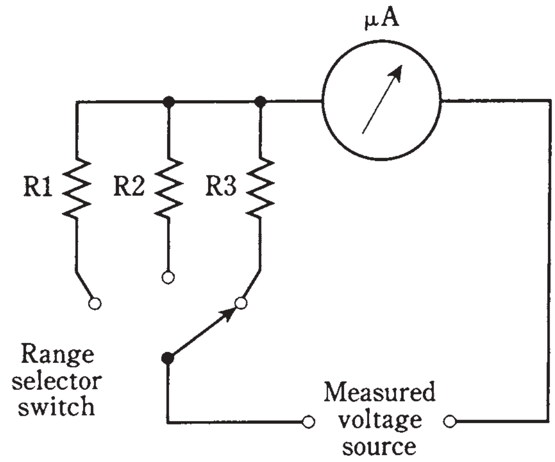
voltage placed across it. Early electrical experimenters recognized that an ammeter could be used to measure voltage, since an ammeter is a form of constant-resistance circuit.

If you connect an ammeter directly across a source of voltage—a battery, say—the meter needle will deflect. 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 0-10 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. This is because all 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 the 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 to get a meaningful indication on the meter. Using a microammeter and a very large value of resistor in series, a voltmeter can be devised that will draw only a little current from the source.

A voltmeter can be made to have different 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.

It's always good when a voltmeter has a high internal resistance. 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, towards working whatever circuit is hooked up to the supply, and not into just 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 the circuit to be affected the instant you connect the voltmeter to the supply. The less current a voltmeter draws, the less it will affect the behavior of anything that is working from the power supply.



3-6 Circuit for using a microammeter to measure voltage.

Another type of voltmeter uses the effect of electrostatic deflection, rather than electromagnetic deflection. You remember that electric fields produce forces, just as do magnetic fields. Therefore, a pair of plates will attract or repel each other if they are charged. The *electrostatic voltmeter* makes use of this effect, taking 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.

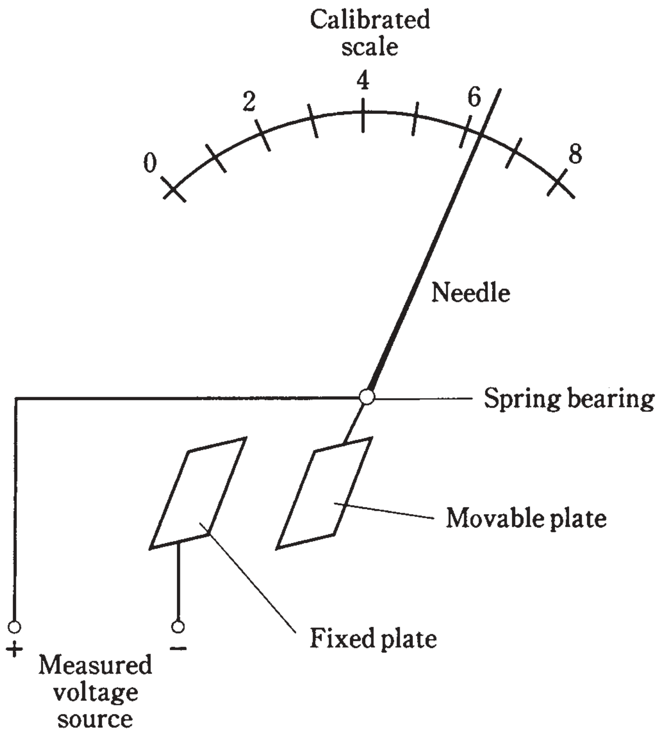
The electrostatic meter draws almost no current from the power supply. The only thing between the plates is air, and air is a nearly perfect insulator. The electrostatic meter will indicate ac as well as dc. The construction tends to be rather delicate, however, and mechanical vibration influences the reading.

Ohmmeters

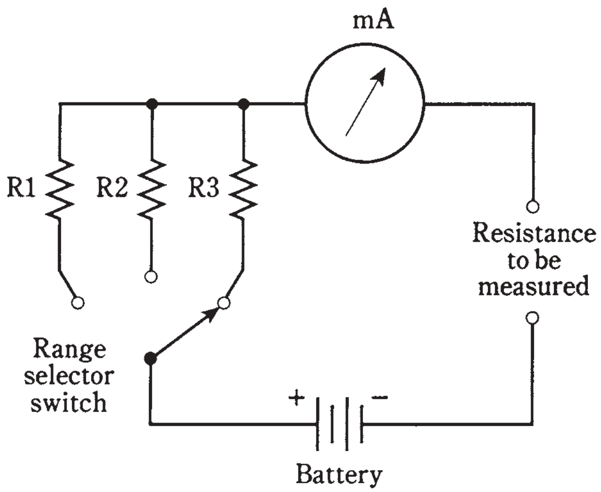
You remember that the current through a circuit depends on the resistance. This principle can be used to manufacture a voltmeter using an ammeter and a resistor. The larger the value of the resistance in series with the meter, the more voltage is needed to produce a reading of full scale. This has a converse, or a “flip side”: Given a constant voltage, the current through the meter will vary if the resistance varies. This provides a means for measuring resistances.

An *ohmmeter* is almost always constructed by means of 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 will give indications in ohms over any desired range. Usually, zero on the meter 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, or 10 KΩ (10,000 Ω).

Ohmmeters must be precalibrated at the factory where they are made. 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. It is also necessary that the battery be exactly the right kind, and that it be reasonably fresh so that it will provide



3-7 Simplified drawing of an electrostatic voltmeter.

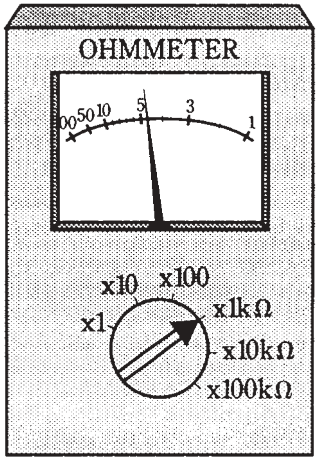


3-8 Circuit for using a milliammeter to measure resistance.

the appropriate voltage. The smallest deviation from the required voltage can cause a big error in the meter indication.

The scale of an ohmmeter is nonlinear. That is, the graduations are not the same everywhere. Values tend to be squashed together towards the "infinity" end of the scale.

It can be difficult to interpolate for high values of resistance, unless the right scale 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 until the meter 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 says 4.7, but the range switch says 1 KΩ. This indicates a resistance of 4.7 KΩ, or 4700 Ω.



3-9 An example of an ohmeter reading. This device shows about 4.7 × 1K = 4.7 K = 4700 ohms.

Ohmmeters will 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 battery voltage. This in effect changes the battery voltage, and the meter reading is thrown way off. Sometimes the meter might even read "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 need to 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 too much 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; larger voltages require special leads 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 resistance is on the order of several megohms or tens of megohms. The lower limit of resistance indication is about an ohm.

FET and vacuum-tube voltmeters

It was mentioned that a good voltmeter will disturb the circuit under test as little as possible, and this requires that the meter have a high internal resistance. Besides the electrostatic type voltmeter, there is another way to get an extremely high internal resistance. This is to sample a tiny, tiny current, far too small for any meter to directly indicate, and then amplify this current so that a meter will show it. When a miniscule amount of current is drawn from a circuit, the equivalent resistance is always extremely high.

The most effective way to accomplish the amplification, while making sure that the current drawn really is tiny, is to use either a *vacuum tube* or a *field-effect transistor (FET).* You needn't worry about how such amplifiers work right now; that subject will come much later in this book. A voltmeter that uses a vacuum tube amplifier to minimize the current drain is known as a *vacuum-tube voltmeter (VTVM).* If an FET is used, the meter is called a *FET voltmeter (FETVM).* Either of these devices provide an extremely high input resistance along with good sensitivity and amplification. And they allow measurement of lower voltages, in general, than electrostatic voltmeters.

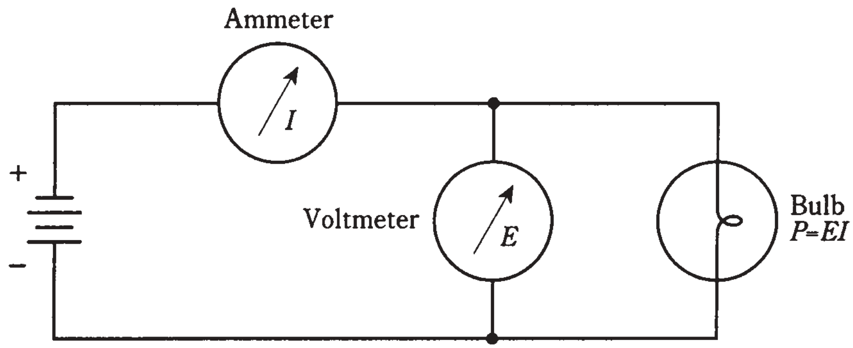
Wattmeters

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

You might think that you can just 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. And in fact, for practically all dc circuits, this is an excellent way to measure power (Fig. 3-10).

Quite often, however, it's simpler than that. 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 can be measured in the same way as with dc: 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 would show a reading of 117 W, because $P = EI = 117 \times 1 = 117$ W. If the meter reading were 300 W, the current would be 300/117 = 2.56 A.

An electric iron might consume 1000 W, or a current of 1000/117 = 8.55 A. And a large heating unit might gobble up 2000 W, requiring a current of 2000/117 = 17. 1 A. This might blow a fuse or breaker, since these devices are often rated for only 15 A. You've probably had an experience where you hooked up too many appliances to a single circuit, blowing the fuse or breaker. The reason was that the appliances, combined, drew too



3-10 Power can be measured with a voltmeter and an ammeter.

much current for the house wiring to safely handle, and the fuse or breaker, detecting the excess current, opened the circuit.

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 as their indicating devices.

Watt-hour meters

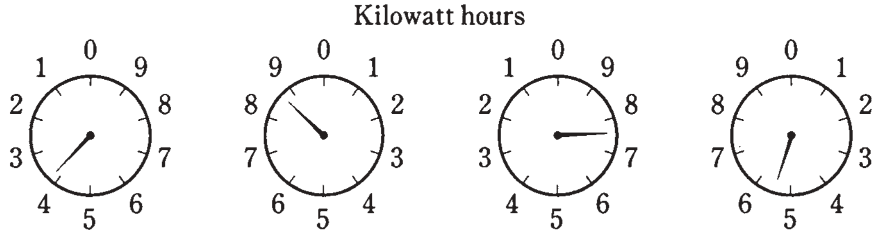
The utility company is not too interested in how much power you're using with one appliance, or even how much power a single household is drawing, at any given time. By far the greater concern is the total energy that is used over a day, a week, a month or a year. Electrical energy is measured in watt hours, or, more commonly for utility purposes, in kilowatt hours (kWh). The device that indicates this is the *watt-hour meter* or *kilowatt-hour meter.*

The most often-used means of measuring electrical energy is by using a small electric motor device, whose speed 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 house, apartment or building. This is usually at a point where the voltage is 234 V. This is split into some circuits with 234 V, for heavy-duty appliances such as the oven, washer and dryer, and the general household fines for lamps, clock radios and, television sets.

You've surely seen the little disk in the utility meter going around and around, sometimes fast, other times slowly. Its speed depends on the power you're using. 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 for electricity.

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 shown in Fig. 3-11. Read from left to right. For each little meter, take down the number that the pointer has most recently passed. Write down the rest as you go. The meter in the figure reads 3875 kWh. If you want to be really precise, you can say it reads 3875-1/2 kWh.



3-11 An example of a utility meter. The reading is a little more than 3875 kWh.

Digital readout meters

Increasingly, metering devices are being designed so that they provide a direct readout, and there's no need (or possibility) for interpolation. The number on the meter is the indication. It's that simple. Such a meter is called a *digital meter.*

The advantage of a digital meter is 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 very well when the value of the quantity does not change very often or very fast.

But 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 would 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 the "feel" of an analog meter, even in situations where a digital meter would work just as well.

One problem you might have 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 an analog meter, too.

Frequency counters

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

The *frequency counter* measures 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 even billions of pulses per second, and it shows the rate on a digital display that is as easy to read as a digital watch. It measures frequency directly by tallying up the number of pulses in an oscillating wave, even when the number of pulses per second is huge.

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 be good by one additional digit. Modern frequency counters are good to six, seven or eight digits; sophisticated lab devices will show frequency to nine or ten digits.

Other specialized meter types

The following are some less common types of meters that you might come across in electrical and electronic work.

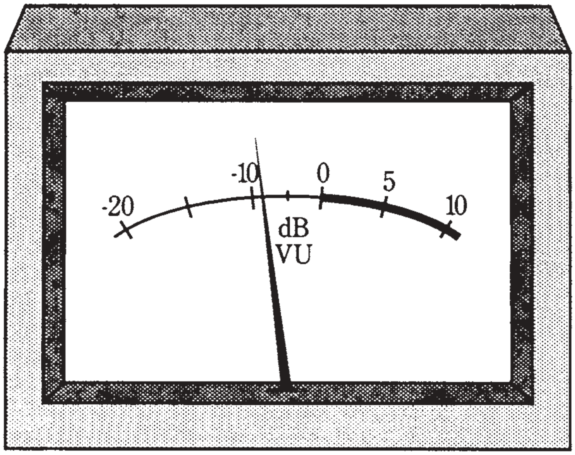
VU and decibel meters

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 sometimes encounter 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.

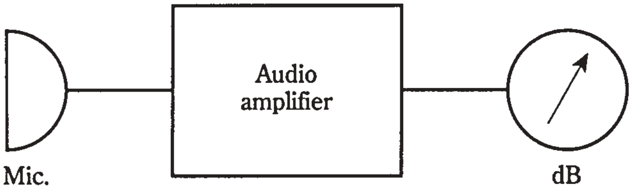
Audio loudness is given in *volume units (VU),* and the meter that indicates it is called a *VU meter.* Usually, such meters have a zero marker with a red line to the right and a black line to the left, and they are calibrated in decibels (dB) above and below this zero marker (Fig. 3-12). The meter might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

As music is played through the system, or as a voice comes over it, the VU meter needle will kick 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 probably 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, standardized sensitivity (Fig. 3-13). You have perhaps heard that a vacuum cleaner will produce 80 dB of sound, and a large truck going by might subject your ears to 90 dB. These figures are determined by a sound-level meter. A VU meter is a special form of sound-level meter.



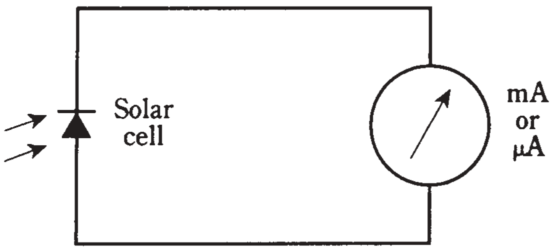
3-12 A VU meter. The heavy scale is usually red, indicating high risk of audio distortion.



3-13 A sound-level meter.

Light meters

Light intensity is measured by means of a light *meter* or *illumination meter.* You might think that it's easy to make this kind of meter by connecting a milliammeter to a solar (photovoltaic) cell. And this is, in fact, a good way to construct an inexpensive light meter (Fig. 3-14). More sophisticated devices might use dc amplifiers to enhance sensitivity and to allow for several different ranges of readings.



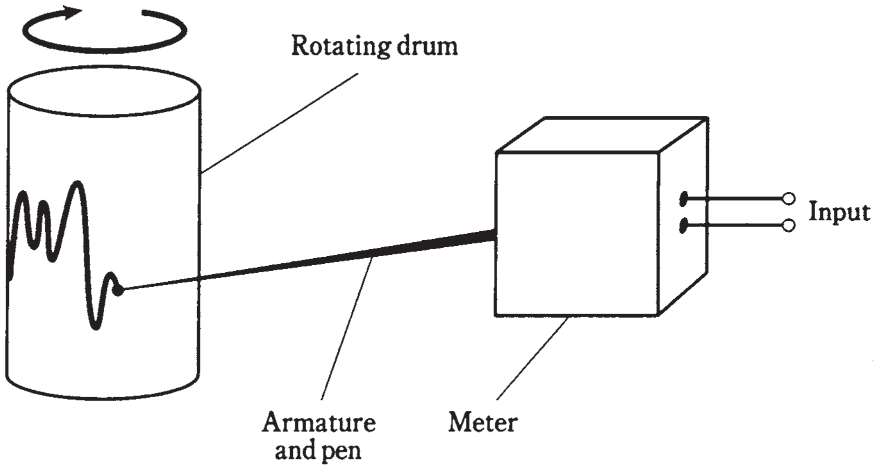
3-14 A simple light meter.

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 units such as *lumens* or *candela.* It's not important that you know the precise definitions of these units in electricity and electronics.

Sometimes, meters such as the one in Fig. 3-14 are used to measure infrared or ultraviolet intensity. Different types of photovoltaic cells have peak sensitivity at different wavelengths. Filters can be used to block out wavelengths that you don't want the meter to detect.

Pen recorders

A meter movement can be equipped with a marking device, usually a pen, 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 Simplified drawing of a pen recorder.

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

Oscilloscopes

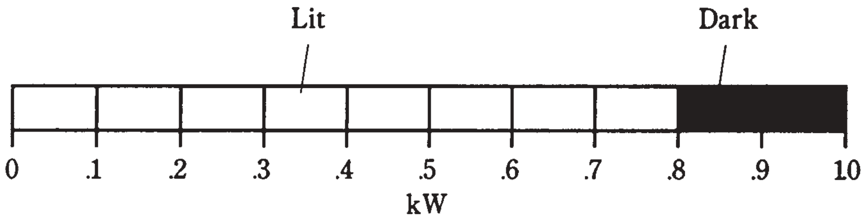
Another graphic meter is the *oscilloscope.* This measures and records quantities that vary rapidly, at rates of hundreds, thousands, or millions of times per second. It creates a "graph" by throwing a beam of electrons at a phosphor screen. A *cathode-ray tube,* similar to the kind in a television set, is employed.

Oscilloscopes are useful for looking at 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 instantaneous 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 kind of "measurement, " since it is qualitative as well as quantitative. If a wave shape "looks wrong," it might indicate distortion in a circuit, or possibly even betray a burned-out component someplace.

Bar-graph meters

A cheap, simple kind of meter can be made using a string of light-emitting diodes (LEDs) or a liquid-crystal display (LCD) along with a digital scale, to indicate approximate levels of current, voltage or power. This type of meter has no moving parts to break, just like a digital meter. But it also 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. It indicates 0.8 kW or 800 watts, approximately.



3-16 A bar-graph meter. This device shows a power level of about 0.8kW or 800W.

The chief disadvantage of the bar-graph meter is that it isn't very accurate. For this reason it is not generally used in laboratory testing. The LED or LCD devices sometimes also flicker when the level is "between" two values given by the bars. This can be annoying to some people.

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 force 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:
   A. Electromagnetic deflection.
   B. Electrostatic force.

C. Magnetic force.

D. Electroscopic force.

3. Suppose a certain current in a galvanometer causes the needle to deflect 20 degrees, and then this current is doubled. The needle deflection:

A. Will decrease.

B. Will stay the same.

C. Will increase.

D. Will reverse direction.

4. One important advantage of an electrostatic meter is that:

A. It measures very small currents.

B. It will handle large currents.

C. It can detect ac voltages.

D. It draws a large current from the source.

5. A thermocouple:

A. Gets warm when current flows through it.

B. Is a thin, straight, special wire.

C. Generates dc when exposed to light.

D. Generates ac when heated.

6. One advantage of an electromagnet meter over a permanent-magnet meter is 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. An ammeter shunt is useful because:

A. It increases meter sensitivity.

B. It makes a meter more physically rugged.

C. It allows for measurement of a wide range of currents.

D. It prevents overheating of the meter.

8. Voltmeters should generally have:

A. Large internal resistance.

B. Low internal resistance.

C. Maximum possible sensitivity.

D. Ability to withstand large currents.

9. To measure power-supply voltage being used by a circuit, a voltmeter

A. Is placed in series with the circuit that works from the supply.

B. Is placed between the negative pole of the supply and the circuit working from the supply.

C. Is placed between the positive pole of the supply and the circuit working from the supply.

D. Is placed in parallel with the circuit that works from the supply.

10. Which of the following will *not* cause a major 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 range switch selection.

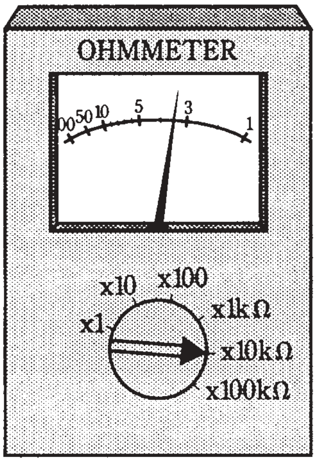
11. The ohmmeter in Fig. 3-17 shows a reading of about:

A. 33,000 Ω.

B. 3.3 KΩ.

C. 330 Ω.

D. 33 Ω.



3-17 Illustration for quiz question 11.

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 as well as to dc.

13. Which of the following is *not* a function of a fuse?

A. To be sure 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 circuit can deliver.

D. To make sure the current is within safe limits.

14. A utility meter's motor speed works directly from:

 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 usually 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 for measurement of:

 A. Sound.

 B. Decibels.

 C. Power.

 D. Energy.

18. The meter movement in an illumination meter measures:

 A. Current.

 B. Voltage.

 C. Power.

 D. Energy.

19. An oscilloscope *cannot* be used to indicate:

 A. Frequency.

 B. Wave shape.

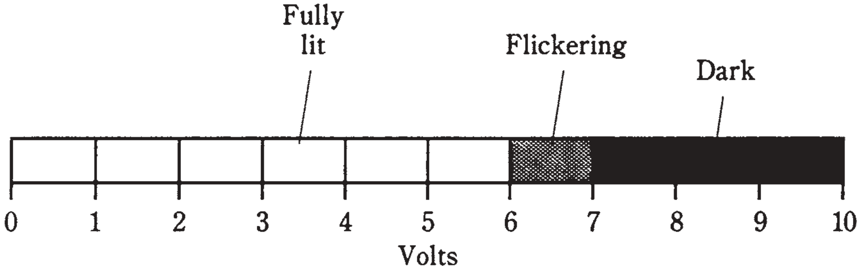
 C. Energy.

 D. Peak signal voltage.

20. The display in Fig. 3-18 could be caused by a voltage of:

 A. 6.0 V.

B. 6.6 V.
C. 7. 0V.
D. No way to tell; the meter is malfunctioning.



3-18 Illustration for quiz question 20.

4
CHAPTER

Basic dc circuits

YOU'VE ALREADY SEEN SOME SIMPLE ELECTRICAL CIRCUIT DIAGRAMS. SOME OF these are the same kinds of diagrams, using the same symbols, that professional technicians and engineers use. 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 direct-current (dc) and low-frequency alternating-current (ac) circuits.

Schematic symbols

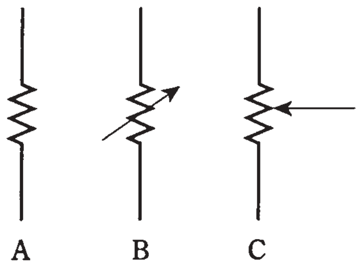
In this course, the plan is to familiarize you with schematic symbols mainly by getting you to read and use them "in action," rather than by dryly drilling you with them. But it's a good idea now to check Appendix B and look over the various symbols. Some of the more common ones are mentioned here.

The simplest schematic symbol is the one representing a wire or electrical conductor: a straight, solid line. Sometimes dotted lines are used to represent conductors, but usually, dotted 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, so that the imaginary charge carriers are forced to march in formation like soldiers. 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.

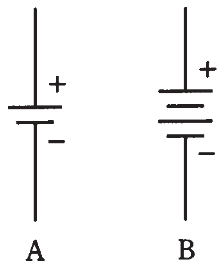
A resistor is indicated by a zig-zaggy line. A variable resistor, or potentiometer, is indicated by a zig-zaggy line with an arrow through it, or by a zig-zaggy line with an arrow pointing at it. These symbols are shown in Fig. 4-1.

A cell is shown by two parallel lines, one longer than the other. The longer line represents the plus terminal. A battery, or combination of cells in series, is indicated by



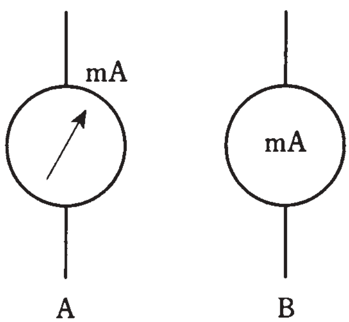
4-1 At A, a fixed resistor. At B, a two-terminal variable resistor. At C, a three terminal potentiometer.

four parallel lines, long-short-long-short. It's not necessary to use more than four lines for any battery, even though sometimes you'll see six or eight lines. The symbols for a cell and a battery are shown in Fig. 4-2.



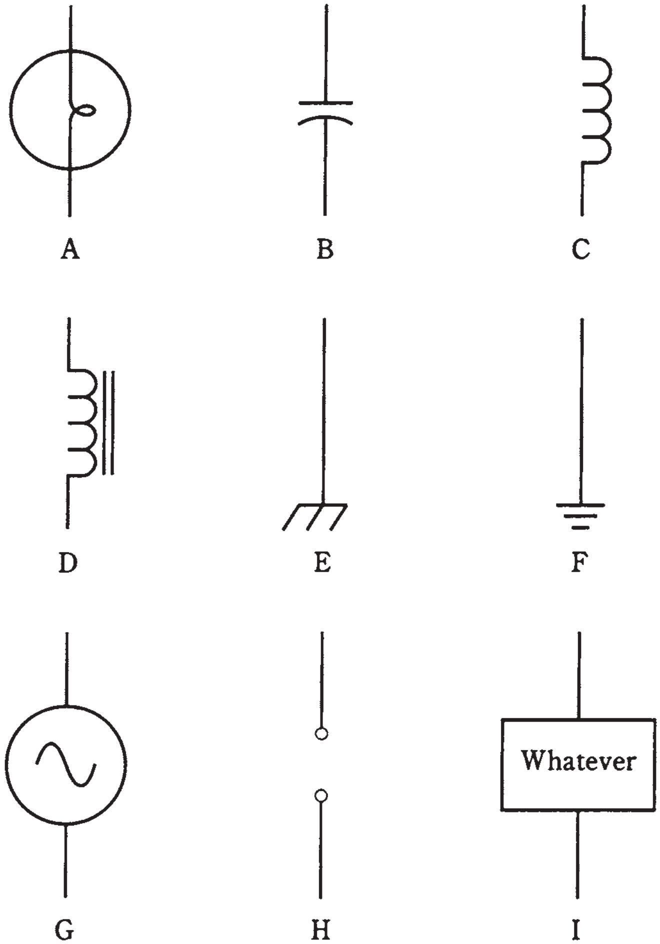
4-2 At A, a single cell. At B, a battery.

Meters are indicated as circles. Sometimes the circle has an arrow inside it, and the meter type, such as mA (milliammeter) or V (voltmeter) are 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 it's done, as long as you're consistent everywhere in a schematic diagram.



4-3 Meter symbols. At A, designator outside; at B, designator inside. Either symbol is OK.

Some other common symbols include the lamp, the capacitor, the air-core coil, the iron-core coil, the chassis ground, the earth ground, the alternating-current source, the set of terminals, and the "black box," a rectangle with the designator written inside. These are shown in Fig. 4-4.



4-4 Nine common schematic symbols. A: Incandescent lamp. B: Capacitor. C: Air-core coil. D: Iron-core coil. E: Chassis ground. F: Earth ground. G: Source of alternating current (ac). H: Pair of terminals. I: Specialized component or device.

Schematic diagrams

Look back through the earlier chapters of this book and observe the *schematic diagrams.* These are all simple examples of how professionals would draw the circuits. There is no inscrutable gobbledygook to put in to make them into the sorts of circuit maps that the most brilliant engineer would need. The diagrams you have worked with are exactly like the ones that the engineer would use to depict these circuits.

Wiring diagrams

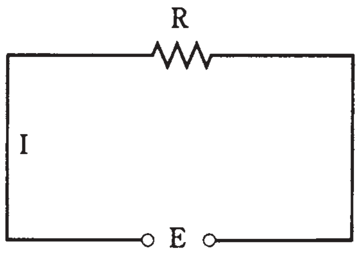
The difference between a schematic diagram and a *wiring diagram* is the amount of detail included. In a schematic diagram, the interconnection of the components is shown, but the actual values of the components are not necessarily indicated.

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.

Suppose you want to build the circuit. You go to an electronics store to get the parts. What sizes 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 them ready made? 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, a jazzed-up schematic. You might have seen this kind of diagram in the back of the instruction manual for a hi-fi amp or an FM stereo tuner or a television set. Wiring diagrams are especially useful and necessary when you must service or repair an electronic device.

Voltage/current/resistance circuit

Most dc circuits can be ultimately boiled down to three major components: a voltage source, a set of conductors, and a resistance. This is shown in the schematic diagram of Fig. 4-5. The voltage or EMF source is called E; the current in the conductor is called I; the resistance is called R. The standard units for these components are the volt, the ampere, and the ohm respectively.



4-5 Simple dc circuit. The voltage is E, the current is I, and 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 also change. If you make the resistance smaller, the current will get larger. If you make the EMF source smaller, the current will decrease. If the current in the circuit increases, the voltage across the resistor will increase. There is a simple arithmetic relationship between these three quantities.

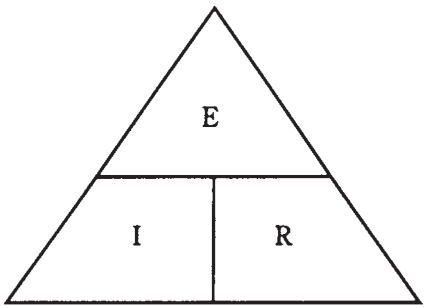
Ohm's Law

The interdependence between current, voltage, and resistance is one of the most fundamental rules, or laws, in electrical circuits. It is called Ohm's Law, named after the scientist who supposedly first expressed it. Three formulas denote this law:

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

You need only to remember the first one in order to derive the others. The easiest way to remember it is to learn the abbreviations E for EMF or voltage, I for current and R for resistance, and then remember that they appear in alphabetical order with the equals sign after the E.

Sometimes the three symbols are written in a triangle, as in Fig. 4-6. To find the value of one, you cover it up and read the positions of the others.



4-6 Ohm's Law triangle.

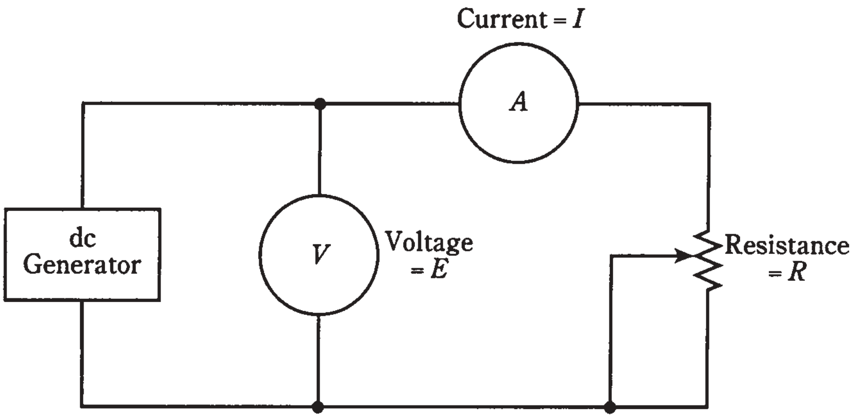
It's important to remember that you must use units of volts, amperes, and ohms in order for Ohm's Law to work right. If you use volts, milliamperes, and ohms or kilovolts, microamperes, and megohms you cannot expect to get the right answers.

If the initial quantities are given in units other than volts, amperes, and ohms, you must convert to these units, then calculate. After that, you can convert the units back again to whatever you like. For example, if you get 13,500,000 ohms as a calculated resistance, you might prefer to say that it's 13.5 megohms.

Current calculations

The first way to use Ohm's Law is to find current values 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 variable dc generator, a voltmeter, some wire, an ammeter, and a calibrated, wide-range potentiometer. Component values have been left out of this diagram, so it's not a wiring diagram. But values can be assigned for the purpose of creating sample Ohm's Law problems. While calculating the current in the following problems, it is necessary to mentally "cover up" the meter.



4-7 Circuit for working Ohm's Law problems.

Problem 4-1

Suppose that the dc generator (Fig. 4-7) produces 10 V, and that the potentiometer is set to a value of 10 Ω. Then what is the current?

This is easily solved by the formula I = E/R. Just plug in the values for E and R; they are both 10, because the units were given in volts and ohms. Then $I = 10/10 = 1$ A.

Problem 4-2

The dc generator (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 K Ω = 10,000 Ω. Then plug the values in: $I = 100/10{,}000 = 0.01$ A. This might better be expressed as 10 mA.

Engineers and technicians prefer to keep the numbers within reason when specifying quantities. Although it's perfectly all right to say that a current is 0.01 A, it's best if the numbers can be kept at 1 or more, but less than 1,000. It is a little silly to talk about a current of 0.003 A, or a resistance of 107,000 Ω, when you can say 3 mA or 107 KΩ.

Problem 4-3

The dc generator (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. The resistance is first changed to ohms, giving 477,000,000 Ω. Then you plug into the Ohm's Law formula: $I = E/R = 88.5/477{,}000{,}000 = 0.000000186$ A = 0.186 uA. This value is less than 1, but there isn't much you can do about it unless you are willing to use units of *nanoamperes (nA),* or billionths of an ampere. Then you can say that the current is 186 nA.

Voltage calculations

The second use of Ohm's Law is to find unknown voltages when the current and the resistance are known. For the following problems, uncover the ammeter and cover the voltmeter scale instead in your mind.

Problem 4-4

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

Use the formula $E = I R$. First, convert the current to amperes: 10 mA = 0. 01 A. Then multiply: $E = 0.01 \times 100 = 1$ V. That's a low, safe voltage, a little less than what is produced by a flashlight cell.

Problem 4-5

Adjust the potentiometer (Fig. 4-7) to a value of 157 KΩ, and let the current reading be 17 mA. What is the voltage of the source?

Now you have to convert both the resistance and the current values to their proper units. A resistance of 157 KΩ is 157,000 Ω; a current of 17 mA is 0. 017 A. Then $E = IR = 0.017 \times 157{,}000 = 2669$ V = 2.669 kV. You might want to round this off to 2.67 kV. This is a dangerous voltage. If you touch the terminals you'll get clobbered.

Problem 4-6

You set the potentiometer (Fig. 4-7) so that the meter reads 1.445 A, and you observe that the potentiometer scale shows 99 ohms. 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 to 143 V. A purist would go further and round it to the nearest 10 volts, to 140 V.

It's never a good idea to specify your answer to a problem with more significant figures than you're given. The best engineers and scientists go by the *rule of significant figures*: keep to the *least* number of digits given in the data. If you follow this rule in Problem 4-6, you must round off the answer to two significant figures, getting 140 V, because the resistance specified (99 Ω) is only accurate to two digits.

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. For the following problems, imagine that both the voltmeter and ammeter scales in Fig. 4-7 are visible, but that the potentiometer is uncalibrated.

Problem 4-7

If the voltmeter reads 24 V and the ammeter shows 3.0 A, what is the value 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 eight and the zero, 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 (meaning some value between $2^1/_2$ A and $3^1/_2$ A), you would only be entitled to express the answer as 8 Ω (somewhere between $7^1/_2$ and $8^1/_2$ Ω). A zero can be a significant figure, just as well as the digits 1 through 9.

Problem 4-8

What is the value of the resistance 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 $R = E/I = 0.229/0.018 = 13\ \Omega$. You're justified in giving your answer to two significant figures, because the current is only given to that many digits.

Problem 4-9

Suppose the ammeter reads 52 uA 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: $R = 2330/0.000052 = 45{,}000{,}000\ \Omega = 45\ \mathrm{M}\Omega$.

Power calculations

You can calculate the power, in watts, in a dc circuit such as that shown in Fig. 4-7, by the formula $P = EI$ or the product of the voltage in volts and the current in amperes. You might not be given the voltage directly, but can calculate it if you know the current and the resistance.

Remember the Ohm's Law formula for obtaining voltage: $E = IR$. If you know I and R, but don't know E, you can get the power P by means of the formula $P = (IR)I = I^2R$. That is, you take the current in amperes, multiply this figure by itself, and then multiply the result by the resistance in ohms.

You can also get the power if you aren't given the current directly. Suppose you're given only the voltage and the resistance. Remember the Ohm's Law formula for obtaining current: $I = E/R$. Therefore, $P = E(E/R) = E^2/R$. Take the voltage, multiply it by itself, and divide by the resistance.

Stated all together, these power formulas are:

$$P = EI = I^2R = E^2/R$$

Now you are ready to do some problems in power calculations. Refer once again to Fig. 4-7.

Problem 4-10

Suppose that the voltmeter 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 zero counts as a significant digit.) Then $P = EI = 12 \times 0.050 = 0.60$ W.

You might say that this is 600 mW, although that is to three significant figures. It's not easy to specify the number 600 to two significant digits without using a means of writing numbers called *scientific notation.* That subject is beyond the scope of this discussion, so for now, you might want to say "600 milliwatts, accurate to two significant figures." (You can probably get away with "600 milliwatts" and nobody will call you on the number of significant digits.)

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 dissipated power?

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 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 a ridiculous state of affairs. An ordinary potentiometer, such as the one you would get at an electronics store, dissipating 22 kW, several times more than a typical household. The voltage must be phenomenal. It's not too hard to figure out that such a voltage would burn out the potentiometer so fast that it would be ruined before the little "Pow!" could even begin to register.

Problem 4-13

Just from curiosity, what is the voltage that would cause so much current to be driven through such a large resistance?

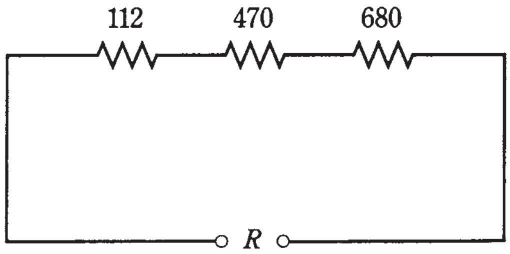
Use Ohm's Law to find the current: $E = IR = 0.680 \times 47{,}000 = 32{,}000$ V $= 32$ kV. That's the sort of voltage you'd expect to find only in certain industrial/commercial applications. The resistance capable of drawing 680 mA from such a voltage would surely not be a potentiometer, but perhaps something like an amplifier tube in a radio broadcast transmitter.

Resistances in series

When you place resistances in series, their ohmic values simply add together to get the total resistance. This is easy to see intuitively, and it's quite simple to remember.

Problem 4-14

Suppose the following resistances are hooked up in series with each other: 112 Ω, 470 Ω, and 680 Ω. What is the total resistance of the series combination (Fig. 4-8)?



4-8 Three resistors in series (Problem 4-14).

Just add 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 components—how precise their actual values are to the ones specified by the manufacturer.

Resistances in parallel

When resistances are placed in parallel, they behave differently than they do in series. In general, if you have a resistor of a certain value and you place other resistors in parallel with it, the overall resistance will decrease.

One way to look at resistances in parallel is to consider them as conductances instead. In parallel, conductances add, just as resistances add 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.

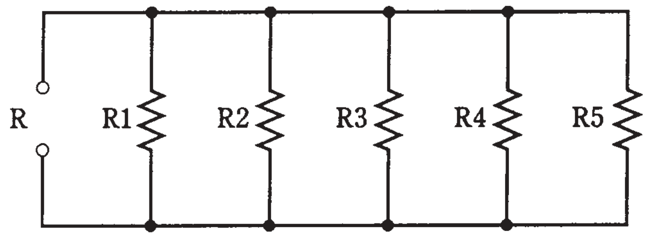
The symbol for conductance is G. This figure, in siemens, is related to the resistance R, in ohms, by the formulas:

$$G = 1/R, \text{ and}$$
$$R = 1/G$$

Problem 4-15

Consider five resistors in parallel. Call them R1 through R5, and call the total resistance R as shown in the diagram Fig. 4-9. Let R1 = 100 Ω, R2 = 200 Ω, R3 = 300 Ω, R4 = 400 Ω and R5 = 500 Ω respectively. What is the total resistance, R, of this parallel combination?

Converting the resistances to conductance values, you get G1 = 1/100 = 0.01 siemens, G2 = 1/200 = 0.005 siemens, G3 = 1/300 = 0.00333 siemens, G4 = 1/400 = 0.0025 siemens, and G5 = 1/500 = 0.002 siemens. Adding these gives G = 0. 01 + 0. 005 + 0.00333 + 0.0025 + 0.002 = 0.0228 siemens. The total resistance is therefore R = 1/G = 1/0.0228 = 43.8 Ω.



4-9 Five resistors in parallel, R1 through R5, give a total resistance R. See Problems 4-15 and 4-16.

When you have resistances in parallel and their values are all equal, the total resistance is equal to the resistance of any one component, divided by the number of components.

Problem 4-16

Suppose there are five resistors R1 through R5 in parallel, as shown in Fig. 4-9, all having a value of 4.7K Ω. What is the total resistance, *R*?

You can probably guess that the total is a little less than 1KΩ or 1000 Ω. So you can convert the value of the single resistor to 4,700 Ω and divide by 5, getting a total resistance of 940 Ω. This is accurate to two significant figures, the 9 and the 4; engineers won't usually be worried about the semantics, and you can just say "940 Ω."

Division of power

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

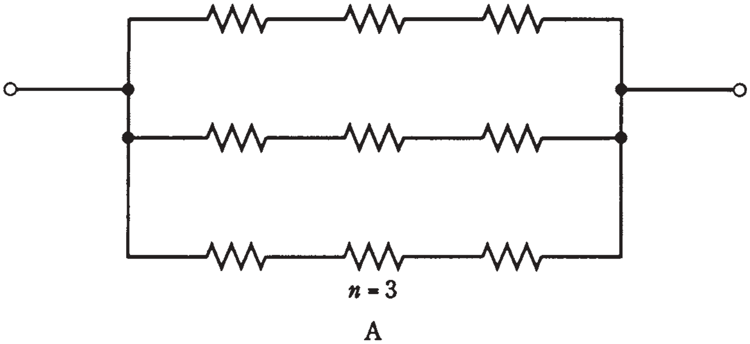
If the resistances in the network all have the same ohmic value, the power from the source will be evenly distributed among the resistances, whether they are hooked up in series or in parallel. If there are eight identical resistors in series with a battery, the network will consume a certain amount 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.

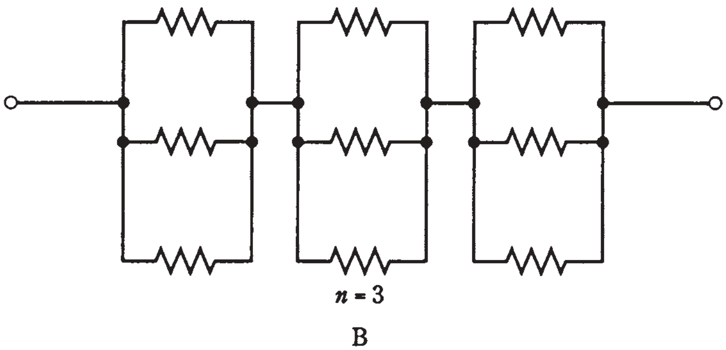
If the resistances in the network do not all have identical ohmic values, they divide up the power unevenly. Situations like this are discussed in the next chapter.

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 (A of Fig. 4-10), or else there are n series sets of n resistors in parallel (B of Fig. 4-10). Either arrangement will give exactly the same results in practice.





4-10 Series-parallel combinations. At A, sets of series resistors are connected in parallel. At B, sets of parallel resistors are in series.

Engineers and technicians sometimes use this to their advantage to get resistors with large power-handling capacity. Each resistor should have the same rating, say 1 W. Then the combination of n by n resistors will have n^2 times that of a single resistor. A 3×3 series-parallel matrix of 2-W resistors can handle $3^2 \times 2 = 9 \times 2 = 18$ W, for example. A 10×10 array of 1-W resistors can take 100 W. Another way to look at this is to see that the total power-handling capacity is multiplied by the total number of individual resistors in the matrix. But this is only true 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. So it's hard to know whether any of the components in such a matrix are going to burn out.

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 alright. 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, will increase the power-handling capability. But in these cases, the total resistance will not be the same as the value of the single resistors. The overall power-handling capacity will always be multiplied by the total number of resistors, whether the network is symmetrical or not, provided all the resistors are the same. In engineering work, cases sometimes do arise where nonsymmetrical networks fit the need just right.

Resistive loads in general

The circuits you've seen here are good for illustrating the principles of dc. But some of the circuits shown here have essentially no practicality. You'll never find a resistor connected across a battery, along with a couple of meters, as shown in Fig. 4-7, for example. The resistor will get warm, maybe even hot, and it will eventually drain the battery in an unspectacular way. Aside from its educational value, the circuit does nothing of any use.

In real life, the ammeter and voltmeter readings in an arrangement such as that shown in Fig. 4-7 would decline with time. Ultimately, you'd be left with a dead, cold battery, a couple of zeroed-out meters, a potentiometer, and some wire.

The resistances in the diagrams like Fig. 4-7 are always put to some use in electrical and electronic circuits. Instead of resistors, you might have light bulbs, appliances (60-Hz utility ac behaves much like dc in many cases), motors, computers, and radios. Voltage division is one important way in which resistors are employed. This, along with more details about current, voltage, and resistance in dc circuits, is discussed in the next chapter.

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 become:
   - A. Four times as great.
   - B. Twice as great.
   - C. The same as it was before.
   - D. Half as great.

2. A wiring diagram would most likely be found in:
   - A. An engineer's general circuit idea notebook.
   - B. An advertisement for an electrical device.
   - C. The service/repair manual for a radio receiver.
   - D. A procedural flowchart.

For questions 3 through 11, see Fig. 4-7.

3. Given a dc voltage source delivering 24 V and a circuit 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 that a circuit has 472 Ω of resistance and the current is 875 mA. Then the source voltage is:
   - A. 413 V.
   - B. 0.539 V.
   - C. 1.85 V.
   - D. None of the above.

5. The dc voltage in a circuit 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 circuit 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. A circuit has a total resistance of 473,332 Ω and draws 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. Then the best expression for the resistance is:
   - A. 15 Ω.
   - B. 15.4 Ω.
   - C. 9.3 Ω.
   - D. 9.32 Ω.

9. 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. The voltage from the source is 12 V and the potentiometer is set for 470 Ω. The power is about:
    - A. 310 mW.
    - B. 25.5 mW.
    - C. 39.2 W.
    - D. 3.26 W.

11. The current through the potentiometer is 17 mA and its value is 1.22KΩ. The power is:
    - 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. Then the total resistance is:
    - A. 90 Ω.
    - B. 3.24 KΩ.
    - C. 540 Ω.
    - D. None of the above.

13. 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 68,000 Ω. Then the total resistance is:

A. 23 Ω.

B. 23 KΩ.

C. 204 Ω.

D. 0.2 MΩ.

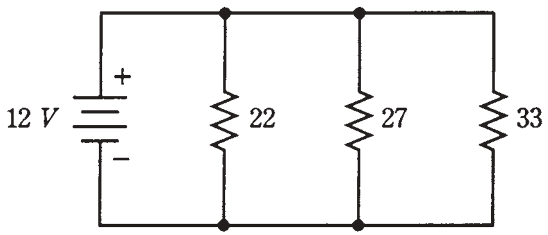
15. There are three resistors in parallel, with values of 22 Ω, 27Ω, and 33 Ω. A 12-V battery is connected across this combination, as shown in Fig. 4-11. What is the current drawn from the battery by this resistance combination?

A. 1.3 A.

B. 15 mA.

C. 150 mA.

D. 1.5 A.



4-11 Illustration for quiz question 15.

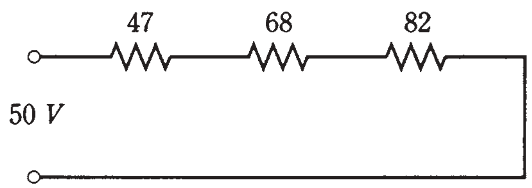
16. Three resistors, with values of 47 Ω, 68 Ω, and 82 Ω, are connected in series with a 50-V dc generator, as shown in Fig. 4-12. The total power consumed by this network of resistors is:

A. 250 mW.

B. 13 mW.

C. 13 W.

D. Not determinable from the data given.



4-12 Illustration for quiz question 16.

17. 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. 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 1000-Ω resistors in series, and connecting these four sets in parallel.

B. Four sets of two 1000-Ω resistors in parallel, and connecting these four sets in series.

C. A 3×3 series-parallel matrix of 1000-Ω resistors.

D. Something other than any of the above.

19. 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 made:

A. From resistors that are all very rugged.

B. From resistors that are all the same.

C. From a series combination of resistors in parallel.

D. From a parallel combination of resistors in series.

A good score is at least 18 correct answers. The answers are in the back of the book.

5 CHAPTER

Direct-current circuit analysis

IN THIS CHAPTER, YOU'LL LEARN MORE ABOUT DC CIRCUITS AND HOW THEY behave. These principles apply to almost all circuits in utility ac applications, too.

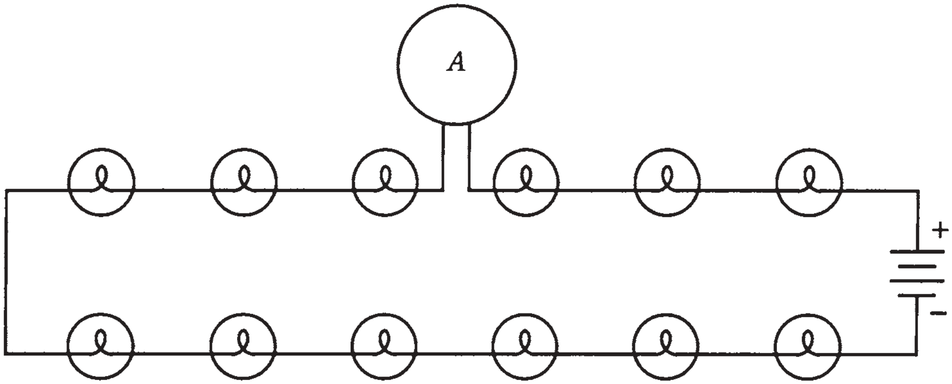
Sometimes it's necessary to analyze networks that don't have obvious practical use. But even a passive network of resistors can serve to set up the conditions for operation of a complex electrical device such as a radio amplifier or a digital calculator, by providing specific voltages or currents.

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. 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 were of 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 would go down. This would force each of the remaining bulbs to carry too much current. Another bulb would probably burn out before long. If it, too, were replaced with a short circuit, the current



5-1 Light bulbs in series. An ammeter, A is placed in the circuit to measure current.

would be increased still further. A third bulb would probably blow out almost right away after the string was plugged in.

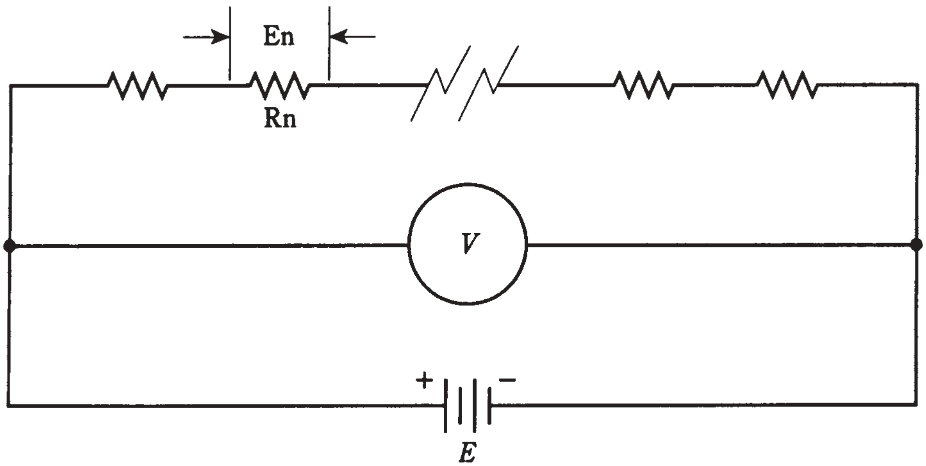
Voltages across series resistances

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 will have a potential difference of 10 V across it. This will be true no matter how large or small the bulbs are, as long as they're all identical.

If you think about this for a moment, it's easy to see why it's true. Look at the schematic diagram of Fig. 5-2. Each resistor carries the same current. Each resistor Rn has a potential difference En across it, equal to the product of the current and the resistance of that particular resistor. These En's are in series, like cells in a battery, so they add together. What if the En's 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" some place, adding or taking away voltage. But there is no such. An EMF cannot come out of nowhere. This principle will be formalized later in this chapter.

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 meter V also shows the sum of the En's across the set of resistors, because it's connected across the set of resistors. The meter says the same thing whether you think of it as measuring the battery voltage E, or as measuring the sum of the En's across the series combination of resistors. Therefore, E is equal to the sum of the En's.

This is a fundamental rule in series dc circuits. It also holds for 60-Hz utility ac circuits almost all the time.



5-2 Analysis of voltage in a series circuit. See text for discussion.

How do you find the voltage across any particular resistor Rn in a circuit like the one in Fig. 5-2? Remember Ohm's Law for finding voltage: $E = IR$. The voltage is equal to the product of the current and the resistance. 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.

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 Vdc. What is the voltage across one of the 10 Ω resistors? Across one of the 20 Ω resistors?

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

$$\text{If Rn} = 10\ \Omega\text{, then En} = I(\text{Rn}) = 0.1 \times 10 = 1.0\ \text{V.}$$
$$\text{If Rn} = 20\ \Omega\text{, then En} = I(\text{Rn}) = 0.1 \times 20 = 2.0\ \text{V.}$$

You can check to see whether 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 V + 10 V = 15 V.

Problem 5-2

In the circuit of Fig. 5-2, what will happen to the voltages across the resistors if one of the 20-Ω resistors is shorted out?

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. This is rounded off to two significant figures.

The voltage En across Rn = 10 Ω is equal to En = I(Rn) = $0.12 \times 10 = 1.2$ V.

The voltage En across Rn = 20 Ω is En = I(Rn) = 0.12 × 20 = 2.4 V.

Checking the total voltage, we add (5 × 1.2) + (4× 2.4) = 6.0 + 9.6 = 15.6 V. This rounds off to 16 V. Where did the extra volt come from?

The above is an example of what can happen when you round off to significant figures and then go through a problem a different way. The rechecking process is not part of the original problem. The answers you got the first time are perfectly alright. The figure 16 V is the result of a kind of mathematical trick, a "gremlin."

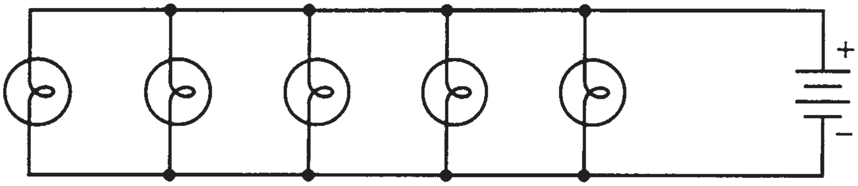
If this phenomenon bothers you, go ahead and keep all the digits your calculator will hold, while you do Problem 5-2 and recheck. The current in the circuit, as obtained by means of a calculator, is 0.115384615 A. When you find the voltages to all these extra digits and recheck, the error will be so tiny that it will cancel itself out, and you'll get a final rounded-off figure of 15 V rather than 16 V.

Some engineers wait until they get the final answer in a problem before they round off to the allowed number of significant digits. This is because the mathematical bugaboo just described can result in large errors, especially in *iterative processes,* involving calculations that are done over and over many times.

You'll probably never be faced with situations like this unless you plan to become an electrical engineer.

Voltage across parallel resistances

Imagine now 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 much easier to fix a parallel-wired string of holiday lights if one bulb should burn out than it is to fix a series-wired string. And the failure of one bulb does not cause catastrophic system failure. In fact, it might be awhile before you notice that the bulb is dark, because all the other ones will stay lit, and their brightness will not change.



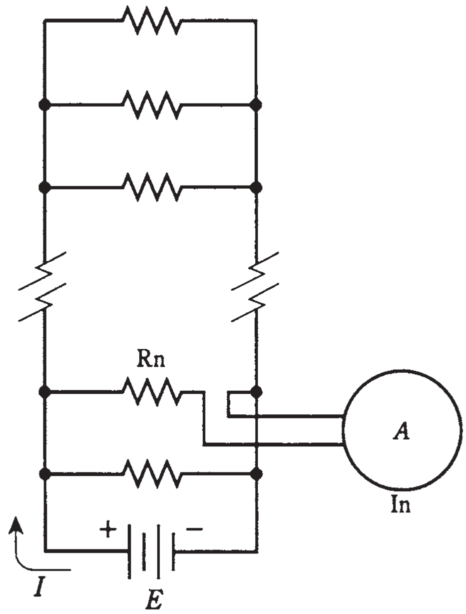
5-3 Light bulbs in parallel.

In a parallel circuit, the voltage across each component is always the same and it is always 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 work independently, as opposed to the series-wired circuit in which they all interact.

If any one branch of a parallel circuit is taken away, the conditions in the other branches will remain the same. If new branches are added, assuming the power supply can handle the load, conditions in previously existing branches will not be affected.

Currents through parallel resistances

Refer to the schematic diagram of Fig. 5-4. The resistors are called Rn. The total parallel resistance in the circuit is *R.* The battery voltage is *E*. The current in branch n, containing resistance *Rn,* is measured by ammeter A and is called In.



5-4 Analysis of a current in a parallel circuit. See text for dicussion-

The sum of all the In's in the circuit is equal to the total current, *I,* drawn from the source. That is, the current is divided up in the parallel circuit, similarly to the way that voltage is divided up in a series circuit.

If you're astute, you'll notice that the direction of current flow in Fig. 5-4 is out from the positive battery terminal. But don't electrons flow out of the minus terminal? Yes—but scientists consider current to flow from plus to minus. This is an example of a mathematical convention or "custom." Such things often outlast their appropriateness. Back in the early days of electrical experimentation, physicists had to choose a direction for the flow of current, and plus-to-minus seemed more logical than minus-to-plus. The exact nature of electric current flow wasn't known then. This notation has not been changed. It was feared that tampering with it would just cause confusion; some people would acknowledge the change while others would not. This might lead to motors running the wrong way, magnets repelling when they should attract, transistors being blown out, and other horrors. Just look at the mess caused by the conflict between Fahrenheit and Celsius temperatures, or between miles and kilometers.

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 $Rn = 120$ by 12 to get $R = 10\ \Omega$. Then the current, I, is found by Ohm's Law: $I = E/R = 12/10 = 1.2$ A.

Problem 5-4

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

This involves finding the current in any given branch. The voltage is 12 V across every branch; $Rn = 120$. Therefore In, the ammeter reading, is found by Ohm's Law: $In = E/Rn = 12/120 = 0.10$ A $= 100$ mA.

All of the *In's* should add to get the total current, I .There are 12 identical branches, each carrying 0.1 A; therefore the sum is $0.1 \times 12 = 1.2$ A. It checks out. And there aren't any problems this time with significant figures.

Those two problems were designed to be easy. Here are two that are a little more involved.

Problem 5-5

Three resistors are in parallel across a battery that supplies $E = 12$ V. The resistances are $R1 = 22\ \Omega$, $R2 = 47\ \Omega$, and $R3 = 68\ \Omega$. These resistors carry currents $I1$, $I2$, and $I3$ respectively. What is the current, $I3$, through $R3$?

This is done by means of Ohm's Law, as if $R3$ were the only resistor in the circuit. There's no need to worry about the parallel combination. The other branches do not affect $I3$. Thus $I3 = E/R3 = 12/68 = 0.18$ A $= 180$ mA.

That problem wasn't hard at all. But it would have seemed that way, had you needlessly calculated the total parallel resistance of $R1$, $R2$, and $R3$.

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 $R1$, $R2$, and $R3$ in parallel, and then calculating I based on R. Another, perhaps easier, way is to find the currents through $R1$, $R2$, and $R3$ individually, and then add them up.

Using the first method, first change the resistances Rn into conductances Gn. This gives $G1 = 1/R1 = 1/22 = 0.04545$ siemens, $G2 = 1/R2 = 1/47 = 0.02128$ siemens, and $G3 = 1/R3 = 1/68 = 0.01471$ siemens. Adding these gives $G = 0.08144$ siemens. 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 $= 980$ mA. Note that extra digits are used throughout the calculation, rounding off only at the end.

Now let's try the other method. Find $I1 = E/R1 = 12/22 = 0.5455$A, $12 = E/R2 = 12/47 = 0.2553$ A, and $13 = E/R3 = 12/68 = 0.1765$ A. Adding these gives $I = I1 + I2 + 3 = 0.\ 5455 + 0.2553 + 0.1765 = 0.9773$ A, rounded off to 0.98 A.

Allowing extra digits during the calculation saved my having to explain away a mathematical artifact. It could save you similar chagrin some day. Doing the problem both ways helped me to be sure I didn't make any mistakes in finding the answer to this problem. It could have the same benefit for you, when the option presents itself.

Power distribution in series circuits

Let's switch back now to series circuits. This is a good exercise: getting used to thinking in different ways and to changing over quickly and often.

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 Pn, based on the formula $Pn = I^2Rn$.

Problem 5-7

Suppose we have a series circuit with a supply of 150 V and three resistors: $R1 = 330$ Ω, $R2 = 680$ Ω, and $R3 = 910$ Ω. What is the power dissipated by $R2$?

You must find the current in the circuit. To do this, calculate the total resistance first. Because the resistors are in series, the total is $R = 330 + 680 + 910 = 1920$ Ω. Then the current is $I = 150/1920 = 0.07813$ A $= 78.1$ mA. The power in $R2$ is $P2 = I^2R2 = 0.07813 \times 0.07813 \times 680 = 4.151$ W. Round this off to two significant digits, because that's all we have in the data, to obtain 4.2 W.

The total power dissipated in a series circuit is equal to the sum of the wattages dissipated in each resistor. In this way, the distribution of power in a series circuit is like the distribution of the voltage.

Problem 5-8

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

The first method is to figure out the power dissipated by each of the three resistors separately, and then add the figures up. The power $P2$ is already known. Let's bring it back to the four significant digits while we calculate: $P2 = 4.151$ W. *R*ecall that the current in the circuit is $I = 0.07813$ A. Then $P1 = 0.07813 \times 0.07813 \times 330 = 2.014$ W, and $P3 = 0.07813 \times 0.07813 \times 910 = 5.555$ W. Adding these gives $P = 2.014 + 4.151 + 5.555 = 11.720$ W. Round this off to 12 W.

The second method is to find the series resistance of all three resistors. This is $R = 1920$ Ω, as found in Problem 5-7. Then $P = I^2R = 0.07813 \times 0.07813 \times 1920 = 11.72$ W, again rounded to 12 W.

You might recognize this as an electrical analog of the distributive law you learned in junior-high-school algebra.

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 Pn, dissipated by resistor Rn, is by using the formula $Pn = E^2/Rn$ where E is the voltage of the supply. Recall that this voltage is the same across every resistor in a parallel circuit.

Problem 5-9

A circuit contains three resistances $R1 = 22$ Ω, $R2 = 47$ Ω, and $R3 = 68$ Ω across a voltage $E = 3.0$ V. Find the power dissipated by each resistor.

First find E^2, because you'll be needing that number often: $E^2 = 3.0 \times 3.0 = 9.0$. Then $P1 = 9.0/22 = 0.4091$ W, $P2 = 9.0/47 = 0.1915$ W, $P3 = 9.0/68 = 0.1324$ W. These can be rounded off to $P1 = 0.41$ W, $P2 = 0.19$ W, and $P3 = 0.13$ W. But remember the values to four places for the next problem.

In a parallel circuit, the total power consumed is equal to the sum of the wattages dissipated by the individual resistances. In this respect, the parallel circuit acts like the series circuit. Power cannot come from nowhere, nor can it vanish. It must all be accounted for.

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 $P1$, $P2$, and $P3$. Let's use the four-significant-digit values for "error reduction insurance." The sum is $P = 0.4091 + 0.1915 + 0.13240 =$ 7330 W. This can be rounded to 0.73 W or 730 mW.

The second method involves finding the resistance R of the parallel combination. You can do this calculation yourself, keeping track for four digits for insurance reasons, getting $R = 12.28\ \Omega$. Then $P = E^2/R = 9.0/12.28 = 0.7329$ W. This can be rounded to 0. 73 W or 730 mW.

In pure mathematics and logic, the results are all deduced from a few simple, intuitively appealing principles called *axioms.* You might already know some of these, such as Euclid's geometry postulates. In electricity and electronics, complex circuit analysis can be made easier if you are acquainted with certain axioms, or *laws.* You've already seen some of these in this chapter. They are:

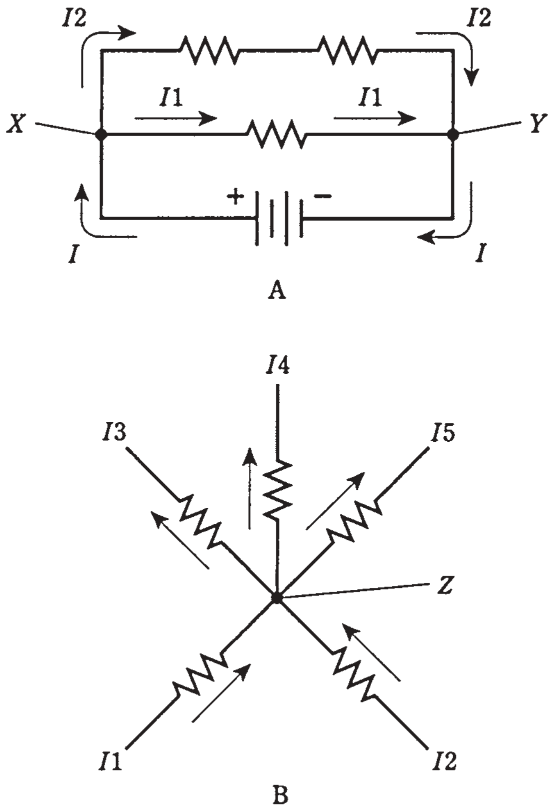
- The current in a series circuit is the same at every point along the way.
- The voltage across any component in a parallel circuit is the same as the voltage across any other, or across the whole set.
- The voltages across elements in a series circuit always add up to the supply voltage.
- The currents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

Now you will learn two of the most famous laws in electricity and electronics. These make it possible to analyze extremely complicated series-parallel networks. That's not what you'll be doing in this course, but given the previous axioms and *Kirchhoff's Laws* that follow, you could if you had to.

Kirchhoff's first law

The physicist Gustav Robert Kirchhoff (1824-1887) was a researcher and experimentalist in electricity back in the time before radio, before electric lighting, and before much was understood about how currents flow.

Kirchhoff reasoned that current must work something like water in a network of pipes, and that the current going into any point has to be the same as the current going out. This is true for any point in a circuit, no matter how many branches lead into or out of the point. Two examples are shown in Fig. 5-5.



5-5 Kirchhoffs First Law. At A, the current into either X or Y is the same as the current out of that point: $I = I1 + I2$. At B, the current into Z equals the current out of Z: $I1 + I2 = I3 + I4 + I5$. Also see quiz questions 13 and 14.

In a network of water pipes that does not leak, and into which no water is added along the way, the total number of cubic feet going in has to be the same as the total volume going out. Water can't form from nothing, nor can it disappear, inside a closed system of pipes. Charge carriers, thought Kirchhoff, must act the same way in an electric circuit.

This is *Kirchhoff's First Law.* An alternative name might be the law of *conservation of current.*

Problem 5-11

Refer to Fig. 5-5A. Suppose all three resistors have values of 100 Ω, and that $I1 = 2.0$ A while $I2 = 1.0$ A. What is the battery voltage?

First, find the current I drawn from the battery. It must be 3.0 A; $I = I1 + I2 = 2.0 + 1.0 = 3.0$ A. Next, find the resistance of the whole combination. The two 100-Ω resistors 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, across the battery, E, is 66.67 Ω. Then $E = IR = 66.67 \times 3.0 = 200$ volts. (Some battery.)

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 Z are 10.0 Ω. The current through each 100-Ω resistor is 500 mA. What is the current through any of the 10.0-Ω resistors, assuming it is equally distributed? What is the voltage, then, across any of the 10. 0-Ω resistors?

The total current into Z is 500 mA + 500 mA = 1.00 A. This must be divided three ways equally among the 10-Ω resistors. Therefore, the current through any one of them is 1.00/3 A = 0.333 A = 333 mA.

The voltage across any one of the 10.0-Ω resistors is 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.

This might sound strange. Surely there is voltage in your electric hair dryer, or radio, or computer. Yes, there is, between different points. But no point can have an EMF with respect to itself. This is so simple that it's almost laughable. A point in a circuit is always shorted out to itself.

What Kirchhoff really was saying, when he wrote his second law, is a more general version of the second and third points previously mentioned. He reasoned that voltage cannot appear out of nowhere, nor can it vanish. All the potential differences must balance out in any circuit, no matter how complicated and no matter how many branches there are.

This is *Kirchhoff's Second Law.* An alternative name might be the law of *conservation of voltage.*

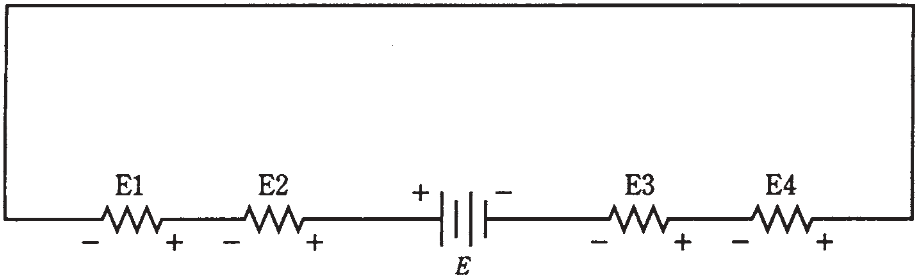
Consider the rule you've already learned about series circuits: The voltages across all the individual resistors add up to the supply voltage. Yes, they do, but the polarities of the EMFs across the resistors are opposite to that of the battery. This is shown in Fig. 5-6. It's a subtle thing. But it becomes clear when a series circuit is drawn with all the components, including the battery or other EMF source, in line with each other, as in Fig. 5-6.

Problem 5-13

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

Find the voltages $E1$, $E2$, $E3$, and $E4$ across each of the resistors. This is done via Ohm's Law. In the case of $E1$, say with the 50-Ω resistor, calculate $E1 = 0.500 \times 50 = 25$ V. In the same way, you can calculate $E2 = 30$ V, $E3 = 35$ V, and $E4 = 40$ V. The supply voltage is the sum $E1 + E2 + E3 + E4 = 25 + 30 + 35 + 40 \text{ V} = 130$ 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 supply in the above example.



5-6 Kirchhoff's Second Law. The sum of the voltages across the resistors is equal to, but has opposite polarity from, the supply voltage E. Thus $E + E1 + E2 + E3 + E4 = 0$. Also see quiz questions 15 and 16.

Problem 5-14

In Fig. 5-6, suppose the battery provides 20 V. Let the resistors, having voltage drops $E1$, $E2$, $E3$, and $E4$, have their ohmic values in the ratio 1:2:3:4 respectively. What is $E3$?

This problem does not tell you the current in the circuit, nor the exact resistance values. But you don't need to know these things. Regardless of what the actual ohmic values are, the ratio $E1{:}E2{:}E3{:}E4$ will be the same. This is a sort of corollary to Kirchhoffs Second Law. You can just invent certain ohmic values with the necessary ratio. Let's have them be $R1 = 1\ \Omega$, $R2 = 2\ \Omega$, $R3 = 3\ \Omega$, and $R4 = 4\ \Omega$. Then the total resistance is $R = R1 + R2 + R3 + R4 = 1 + 2 + 3 + 4 = 10\ \Omega$. You can calculate the current as $I = E/R = 20/10 = 2$ A. Then the voltage $E3$, across $R3$, is given by Ohm's Law as $E3 = I(R3) = 2 \times 3 = 6$ V.

You are encouraged to calculate the other voltages and observe that they add up to 20 V.

In this problem, there is freedom to literally pick numbers out of the air so that calculations are easy. You could have chosen ohmic values like 47, 94, 141, and 188 Ω (these too are in the ratio 1:2:3:4), and you'd still get $E3 = 6$ V. (Go ahead and try it.) But that would have made needless work for yourself.

Series combinations of resistors are often used by electronic engineers to obtain various voltage ratios, to make circuits work just right. These resistance circuits are called *voltage divider networks.*

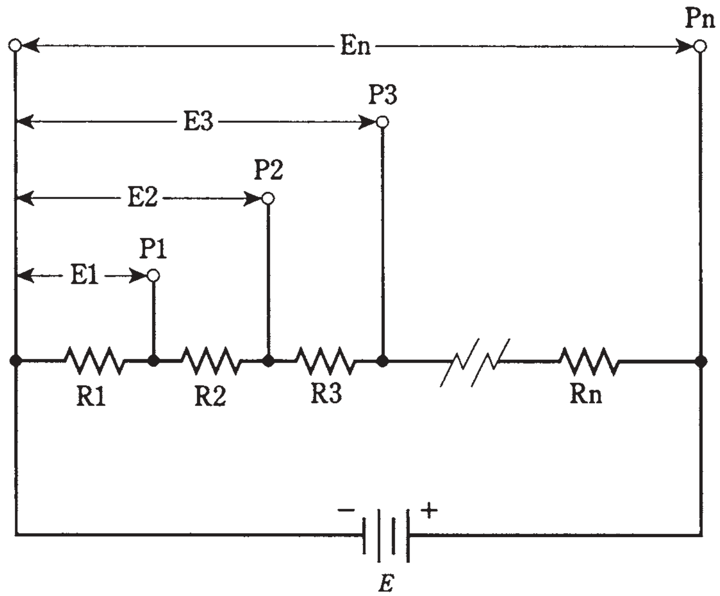
Voltage divider networks

Earlier, you were assured that this course would not drag you through ridiculously complicated circuits. You can imagine, by now nightmarish series-parallel matrixes of resistors drawn all over whole sheets of paper, captioned with wicked queries: "What is the current through R135?" But that stuff is best left to professional engineers, and even they aren't likely to come across it very often. Their job is to make things as neat and efficient as possible. If an engineer actually is faced with such a scenario, the reaction will probably be, "How can this circuit be simplified?"

Resistances in series produce ratios of voltages, and these ratios can be tailored to meet certain needs.

When designing voltage divider networks, the resistance values should be as small as possible, without causing too much current drain on the 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 $R1$, $R2$, $R3$, … Rn. The total resistance is $R = R1 + R2 + R3 + \ldots + Rn$. The supply voltage is E, and the current in the circuit is therefore $I = E/R$. At the various points $P1$, $P2$, $P3$,, … Pn, voltages will be $E1$, $E2$, $E3$, …, En. The last voltage, En, is the same as the supply voltage, E. All the other voltages are less than E, so $E1 < E2 < E3 < \ldots < En = E$. (The symbol < means "is less than.")



5-7 General arrangement for voltage divider circuit. Designators are discussed in the text. Also see quiz questions 19 and 20.

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. The voltage $E1$ is equal to $E \times R1/R$. The voltage $E2$ is equal to $E \times (R1 + R2)/R$. The voltage $E3$ is $E \times (R1 + R2 + R3)/R$. You can mentally continue this process to get each one of the voltages at points all the way up to $En = E \times (R1 + R2 + R3 + \ldots + Rn)/R = E \times R/R = E \times 1 = E$.

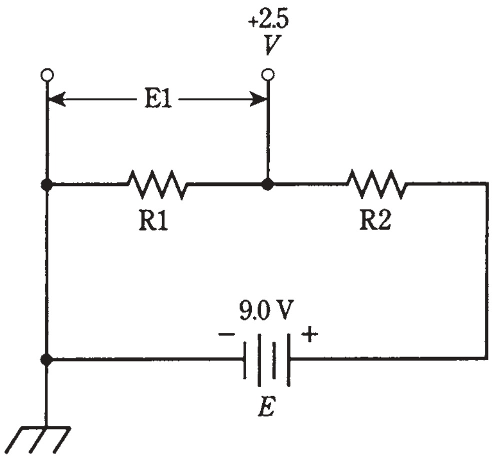
Usually there are only two or three intermediate voltages in a voltage-divider network. So designing such a circuit isn't as complicated as the above formulas might lead you to think.

The following problems are similar to those encountered by electronic engineers.

Problem 5-15

In a transistorized amplifier, the battery supplies 9.0 V. The minus terminal is at common (chassis) ground. At some point, you need to get +2.5 V. Give an example of a pair of resistors that can be connected in series, such that +2.5 V will be provided at some point.

See the schematic diagram of Fig. 5-8. There are infinitely many different combinations of resistances that will work here. You pick some total value, say $R = R1 + R2 = 1000\ \Omega$. You know that the ratio $R1{:}R$ will always be the same as the ratio $E1{:}E$. In this case $E1 = 2.5$ V, so $E1{:}E = 2.5/9.0 = 0.28$. Therefore $R1{:}R$ should be 0.28. Because $R = 1000\ \Omega$, this means $R1 = 280\ \Omega$. The value of $R2$ will be the difference $1000 - 280\ \Omega = 720\ \Omega$.



5-8 Example of a voltage divider network.

Problem 5-16

What is the current drawn by the resistances in the previous problem?

Simply use Ohm's Law to get $I = E/R = 9.0/1000 = 9.0$ mA.

Problem 5-17

Suppose that it is permissible to draw up to 100 MA of current in the problem shown by Fig. 5-8. You, the engineer, want to design the circuit to draw this maximum current, because that will offer the best voltage regulation for the circuit to be operated from the network. What values of resistances $R1$ and $R2$ should you use?

Calculate the total resistance first, using Ohm's Law: $R = E/I = 9.0/0.1 = 90\ \Omega$. The ratio desired is $R1{:}R2 = 2.5/9.\ 0 = 0.28$. Then you would use $R1 = 0.28 \text{ x } 90 = 25\ \Omega$. The value of $R2$ is the remainder: $R2 = 90 - 25 = 65\ \Omega$.

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 holiday ornament bulbs, if one bulb gets shorted out, which of these is most likely?

   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.

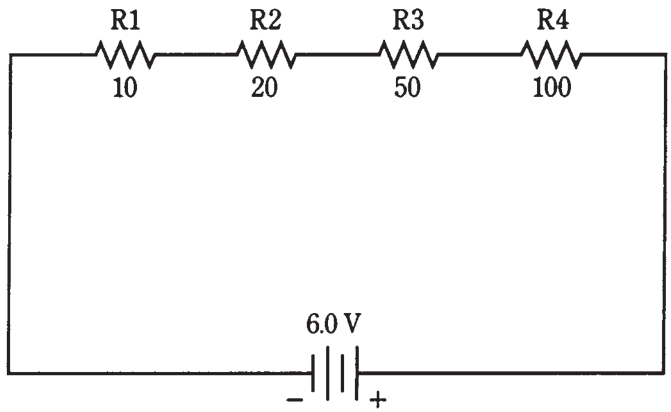
2. Four resistors are connected in series across a 6.0-V battery. The values are $R1 = 10\ \Omega$, $R2 = 20\ \Omega$, $R3 = 50\ \Omega$, and $R4 = 100\ \Omega$ as shown in Fig. 5-9. The voltage across R2 is:

   A. 0.18 V.

   B. 33 mV.

   C. 5.6 mV.

   D. 670 mV.



5-9 Illustration for quiz questions 2, 3, 8, and 9.

3. In question 2 (Fig. 5-9), the voltage across the combination of $R3$ and $R4$ is:

   A. 0.22 V.

   B. 0.22 mV.

   C. 5.0 V.

   D. 3.3 V.

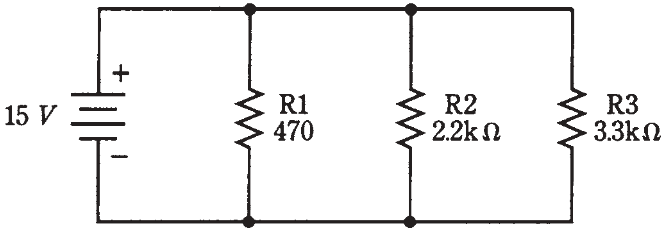
4. Three resistors are connected in parallel across a battery that delivers 15 V. The values are $R1 = 470\ \Omega$, $R2 = 2.2\ \mathrm{K}\Omega$, $R3 = 3.3\ \mathrm{K}\Omega$ (Fig. 5-10). The voltage across $R2$ is:

   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.

5. In the example of question 4 (Fig. 5-10), what is the current through *R*2?

 A. 6.8 mA.
 B. 43 mA.
 C. 150 mA.
 D. 6.8 A.

6. In the example of question 4 (Fig. 5-10), what is the total current drawn from the source?

 A. 6.8 mA.
 B. 43 mA.
 C. 150 mA.
 D. 6.8 A.

7. In the example of question 4 (Fig. 5-10), suppose that resistor R2 opens up. The current through the other two resistors will:

 A. Increase.
 B. Decrease.
 C. Drop to zero.
 D. No change.

8. Four resistors are connected in series with a 6.0-V supply, with values shown in Fig. 5-9 (the same as question 2). What is the power dissipated by the whole combination?

 A. 200 mW.
 B. 6.5 mW.
 C. 200 W.
 D. 6.5 W.

9. In 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. Three resistors are in parallel in the same configuration and with the same values as in problem 4 (Fig. 5-10). What is the power dissipated by the whole set?

A. 5.4 W.

B. 5.4 uW.

C. 650 W.

D. 650 mW.

11. In Fig. 5-10, the power dissipated by $R1$ is:

A. 32 mW.

B. 480 mW.

C. 2.1 W.

D. 31 W.

12. Fill in the blank in the following sentence. In either series or a parallel circuit, the sum of the ____s in each component is equal to the total provided by the supply.

A. Current.

B. Voltage.

C. Wattage.

D. Resistance.

13. Refer to Fig. 5-5A. Suppose the resistors each have values of 33 **Ω**. The battery provides 24 V. The current I1 is:

A. 1.1 A.

B. 730 mA.

C. 360 mA.

D. Not determinable from the information given.

14. Refer to Fig. 5-5B. Let each resistor have a value of 820 **Ω**. Suppose the top three resistors all lead to light bulbs of the exact same wattage. If $I1 = 50$ mA and $I2 = 70$ mA, what is the power dissipated in the resistor carrying current I4?

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 $R1$, $R2$, $R3$, and $R4$ are in the ratio 1:2:4:8 from left to right, and the battery supplies 30 V. Then the voltage $E2$ is:

A. 4 V.
B. 8 V.
C. 16 V.
D. Not determinable from the data given.

16. Refer to Fig. 5-6. Let the resistances each be 3.3 KΩ and the battery 12 V. If the plus terminal of a dc voltmeter is placed between $R1$ and $R2$ (with voltages $E1$ and $E2$), and the minus terminal of the voltmeter is placed between $R3$ and $R4$ (with voltages $E3$ and $E4$), what will the meter register?

A. 0 V.
B. 3 V.
C. 6 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. The battery E is 18.0 V. Suppose there are four resistors in the network: $R1 = 100\ \Omega$, $R2 = 22.0\ \Omega$, $R3 = 33.0\ \Omega$, $R4 = 47.0\ \Omega$. The voltage E3 at P3 is:

A. 4.19 V.
B. 13.8 V.
C. 1.61 V.
D. 2.94 V.

20. Refer to Fig. 5-7. The battery is 12 V; you want intermediate voltages of 3.0,6.0 and 9.0 V. Suppose that a maximum of 200 mA is allowed through the network. What values should the resistors, $R1$, $R2$, $R3$, and $R4$ have, respectively?

A. 15 Ω, 30 Ω, 45 Ω, 60 Ω.
B. 60 Ω, 45 Ω, 30 Ω, 15 Ω.
C. 15 Ω, 15 Ω, 15 Ω, 15 Ω.
D. There isn't enough information to design the circuit.

A good score is at least 18 correct answers. The answers are in the back of the book.

6

CHAPTER

Resistors

AS YOU'VE ALREADY SEEN, ANY ELECTRICAL DEVICE HAS SOME RESISTANCE; none is a perfect conductor. You've also seen some examples of circuits containing components designed to oppose the flow of current. This chapter more closely examines resistors—devices that oppose, control, or limit electrical current.

Why, you might ask, would anyone want to put things into a circuit to reduce the current? Isn't it true that resistors always dissipate some power as heat, and that this invariably means that a circuit becomes less efficient than it would be without the resistor? Well, it's true that resistors always dissipate some power as heat. But resistors can optimize the ability of a circuit to generate or amplify a signal, making the circuit maximally efficient at whatever it is designed to do.

Purpose of the resistor

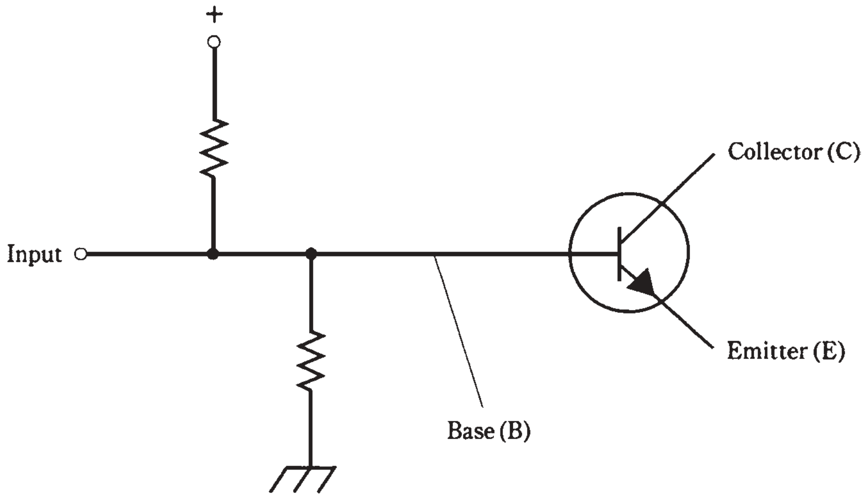
Resistors can play any of numerous different roles in electrical and electronic equipment. Here are a few of the more common ways resistors are used.

Voltage division

You've already learned a little about how voltage dividers can be designed using resistors. The resistors dissipate some power in doing this job, but the resulting voltages are needed for the proper *biasing* of electronic transistors or vacuum tubes. This ensures that an amplifier or oscillator will do its job in the most efficient, reliable possible way.

Biasing

In order to work efficiently, transistors or tubes need the right bias. This means that the control electrode—the *base*, *gate*, or *grid*—must have a certain voltage or current. Networks of resistors accomplish this. Different bias levels are needed for different types of circuits. A radio transmitting amplifier would usually be 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 transistor whose base is biased using a pair of resistors in a voltage-dividing configuration.



6-1 Voltage divider for biasing the base of a transistor.

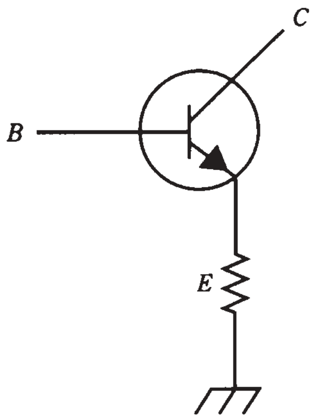
Current limiting

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 in a receiving amplifier. A resistor can keep the transistor from using up a lot of power just getting hot. Without resistors to limit or control the current, the transistor might be overstressed carrying direct current that doesn't contribute to the signal. An improperly designed amplifier might need to have its transistor replaced often, because a resistor wasn't included in the design where it was needed, or because the resistor isn't the right size. Figure 6-2 shows a current-limiting resistor connected in series with a transistor. Usually it is in the emitter circuit as shown in this diagram, but it can also be in the collector circuit.

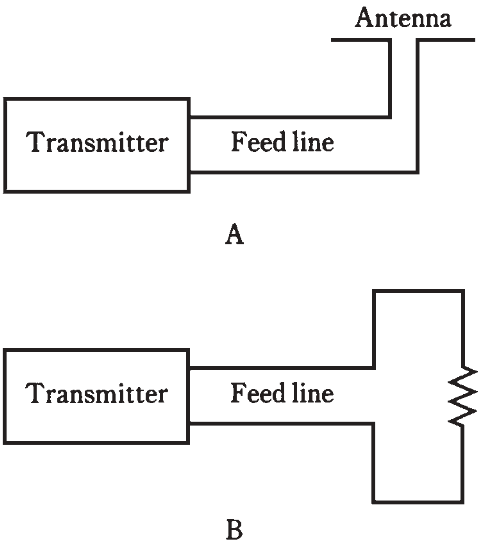
Power dissipation

Dissipating power as heat is not always bad. Sometimes a resistor can be used as a "dummy" component, so that a circuit "sees" the resistor as if it were something more complicated. In radio, for example, a resistor can be used to take the place of an antenna. A transmitter can then be tested in such a way that it doesn't interfere with signals on the airwaves. The transmitter output heats the resistor, without radiating any signal. But as far as the transmitter "knows," it's hooked up to a real antenna (Fig. 6-3).

Another case in which power dissipation is useful is at the input of a power amplifier. Sometimes the circuit *driving* the amplifier (supplying its input signal) has too much



6-2 Current-limiting resistor for a transistor.

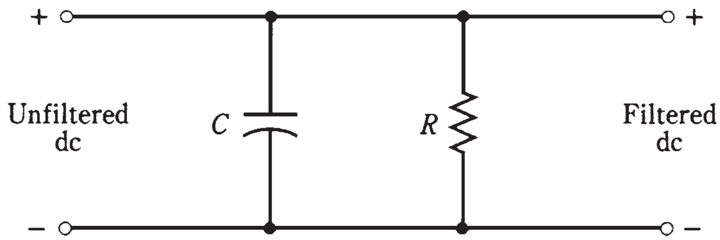


6-3 At A, a transmitter is hooked up to a real antenna; at B, to a resistive "dummy" antenna.

power for the amplifier input. A resistor, or network of resistors, can dissipate this excess so that the power amplifier doesn't get too much drive.

Bleeding off charge

In a high-voltage, direct-current (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 awhile. 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 might get clobbered by this voltage. *Bleeder resistors,* connected across the filter capacitors, drain their stored charge so that servicing the supply is not dangerous (Fig. 6-4).



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

It's always a good idea to short out all filter capacitors, using a screwdriver with an insulated handle, before working on a high-voltage dc power supply. I recall an instance when I was repairing the supply for a radio power amplifier. The capacitors were holding about 2 kV. My supervisor, not very well acquainted with electronics, was looking over my shoulder. I said, "Gonna be a little pop, now," and took a Phillips screwdriver, making sure I had hold of the insulated handle only, and shorted the filter capacitor to the chassis. Bang! I gave my supervisor a brief explanation while he took some deep breaths.

Even if a supply has bleeder resistors, they take awhile to get rid of the residual charge. For safety, always do what I did, whether your supervisor is around or not.

Impedance matching

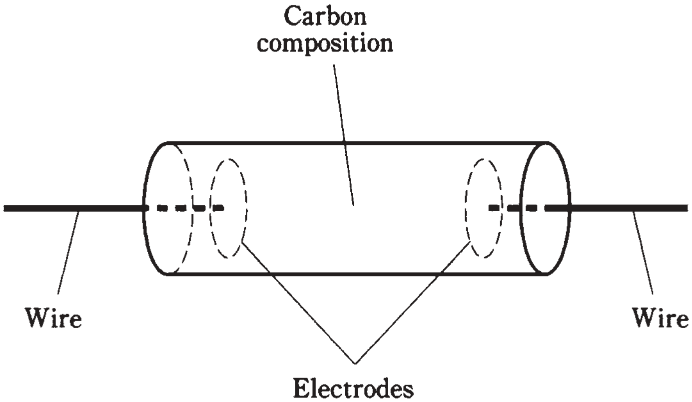
A more subtle, more sophisticated use for resistors is in the *coupling* in a chain of amplifiers, or in the input and output circuits of amplifiers. In order to produce the greatest possible amplification, the *impedances* must agree between the output of a given amplifier and the input of the next.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

Impedance is the alternating-current (ac) cousin of resistance in direct-current (dc) circuits. This is discussed in the next section of this book.

The carbon-composition resistor

Probably the cheapest method of making a resistor is to mix up finely powdered carbon (a fair electrical conductor) with some nonconductive substance, press the resulting clay-like stuff into a cylindrical shape, and insert wire leads in the ends (Fig. 6-5). The resistance of the final product will depend on the ratio of carbon to the nonconducting material, and also on the physical distance between the wire leads. The nonconductive material is usually phenolic, similar to plastic. This results in a carbon-composition resistor.

Carbon-composition resistors can be made to have quite low resistances, all the way up to extremely high resistances. This kind of resistor has the advantage of being pretty much *nonreactive.* That means that it introduces almost pure resistance into the circuit, and not much capacitance or inductance. This makes the *carbon-composition* resistor useful in radio receivers and transmitters.

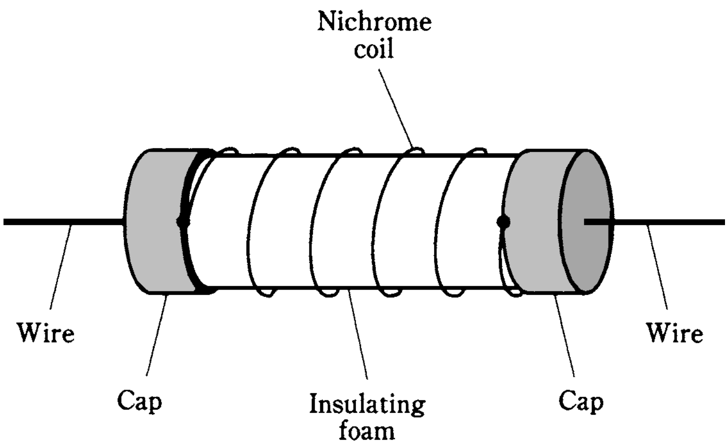


6-5 Construction of a carbon-composition resistor.

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 ¼ W or ½ W. There are ⅛-W units for miniaturized, low- power circuitry, and 1-W or 2-W components for circuits where some electrical ruggedness is needed. Occasionally you'll see a much larger unit, but these are rare.

The wirewound resistor

A more obvious way to get resistance is to use a length of wire that isn't a good conductor. Nichrome is most often used for this. The wire can be wound around a cylindrical form, like a coil (Fig. 6-6). The resistance is determined by how well the wire metal conducts, by its diameter or *gauge,* and by its length. This component is called a *wirewound resistor*.



6-6 Construction of a wirewound resistor.

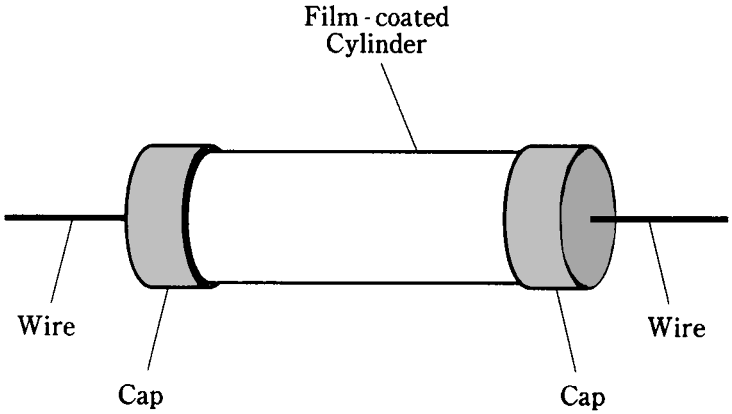
One of the advantages of wirewound resistors is that they can be made to have values within a very close range; that is, they are precision components. Another advantage is that wirewound resistors can be made to handle large amounts of power. Some wirewounds might actually do well as electric heaters, dissipating hundreds, or even thousands of watts.

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, nichrome, or some mixture of ceramic and metal *(cermet)* can be applied to a cylindrical form as a film, or thin layer, in order to obtain a desired value of resistance. This type of resistor is called a *carbon-film resistor* or *metal-film resistor.* It looks like a carbon-composition type, but the construction technique 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 units 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 units, do not have much inductance or capacitance. A disadvantage, in some applications, is that they can't handle as much power as the more massive carbon-composition units or as wirewound types.

Integrated- circuit resistors

Increasingly, whole electronic circuits are being fabricated on semiconductor wafers known as *integrated circuits (ICs)*. It is possible nowadays to put a whole radio receiver into a couple of ICs, or *chips*, whose total volume is about the same as that of the tip of your little finger. In 1930, a similar receiver would have been as large as a television set.

Resistors can be fabricated onto the semiconductor chip that makes up an IC. The thickness, and the types and concentrations of impurities added, control the resistance of the component.

IC resistors can only handle a tiny amount of power because of their small size. But because IC circuits in general are designed to consume minimal power, this is not a problem. The small signals produced by ICs can be amplified using circuits made from discrete components if it is necessary to obtain higher signal power.

The potentiometer

All of the resistors mentioned are fixed in value. It is impossible to change or adjust their resistances. Of course, their values will change if they overheat, or if you chip pieces of them out, but they're meant to provide an unchanging opposition to the flow of electric current.

It might have occurred to you that a variable resistor can be made by hooking up a bunch of fixed resistors in series or parallel, and then switching more or fewer of them in and out. This is almost never done in electronic circuits because there's a better way to get a variable resistance: use a *potentiometer.*

The construction of a potentiometer is shown in simplified form in Fig. 6-8. A resistive strip 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 this 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 graphic equalizers, as the volume controls in some stereo amplifiers, and in some other applications when a linear scale is preferable to a circular scale.

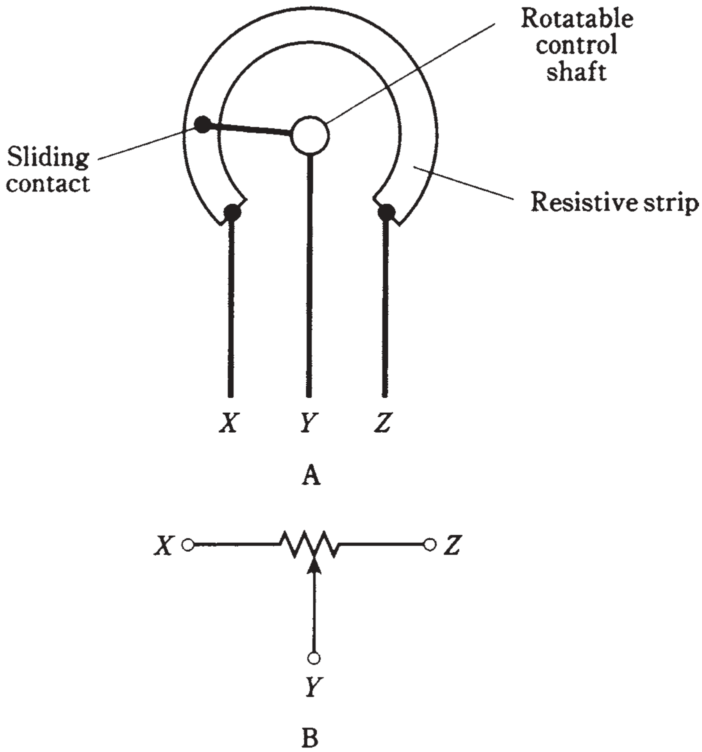
Potentiometers are made to handle only very low levels of current, at low voltage.

Linear taper

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.

Suppose a linear taper potentiometer has a value of zero to 280 Ω. In most units the shaft rotates about 280 degrees, or a little more than three-quarters of a circle. Then the resistance between the center and one end terminal will increase right along with the number of 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. Engineers say that the resistance is a *linear function* of the shaft position.

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



6-8 At A, simplified drawing of the construction of a rotary potentiometer. At B, schematic symbol.

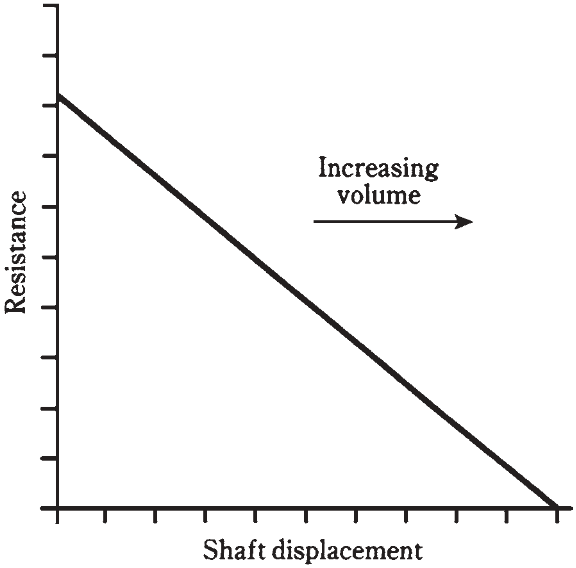
Audio or logarithmic taper

There are some applications for which linear taper potentiometers don't work well. The volume control of a radio receiver is a good example. Your ear/brain perceives sound level according to the *logarithm* of its true level. If you use a linear taper potentiometer as the volume control of a transistor radio or other sound system, the level will seem to go up too slowly in some parts of the control range and too fast in other parts of the control range.

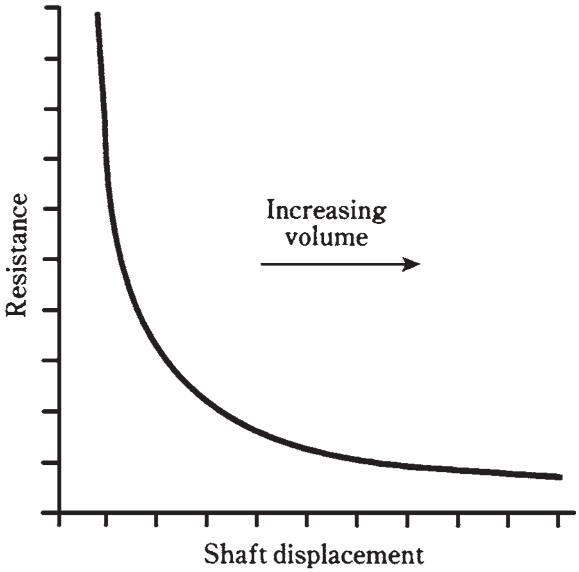
To compensate for the way in which people perceive sound level, an *audio taper* potentiometer is used. In this device, the resistance between the center and end terminal increases in a nonlinear way. This type of potentiometer is sometimes called a *logarithmic-taper* device.

If the shaft is all the way counterclockwise, the volume at the speaker is zero or near zero. If you turn the shaft 30 degrees clockwise, the volume increases to some perceived level; call it one sound unit. If you then turn the volume 30 degrees further clockwise, the volume will seem to go up to two sound units. But in fact it has increased much more than this, in terms of actual sound power.

You perceive sound not as a direct function of the true volume, but in units that are based on the logarithm of the intensity. Audio-taper potentiometers are manufactured so that as you turn the shaft, the sound seems to increase in a smooth, natural way. A graph of resistance versus shaft displacement for an audio-taper potentiometer is shown in Fig. 6-10.



6-9 Resistance-vs-displacement curve for linear taper potentiometer.



6-10 Resistance-vs-displacement curve for audio-taper potentiometer.

This is a good time to sidetrack for a moment and exarnine how sound sensation is measured.

The decibel

Perceived levels of sound, and of other phenomena such as light and radio signals, change according to the logarithm of the actual power level. Units have been invented to take this into account.

The fundamental unit of sound change is called the *decibel*, abbreviated *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 positive decibel values; decreases in volume are negative values.

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

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, and then multiplying the result by 10:

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

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: a 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 change of plus or minus 10 dB is an increase or decrease in sound power of 10 times. A change of plus or minus 20 dB is a hundredfold increase or decrease in sound power. It is not unusual to encounter sounds that range in loudness over plus/minus 60 dB or more—a millionfold variation.

Sound power in terms of decibels

The above formula can be worked inside-out, so that you can determine the final sound power, given the initial sound power and the decibel change.

Suppose the initial sound power is P, and the change in decibels is dB. Let Q be the final sound power. Then Q = P antilog (dB/10).

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

Decibels in real life

A typical volume control potentiometer might have a resistance range such that you can adjust the level over about plus/minus 80 dB. The audio taper ensures that the decibel increase or decrease is a straightforward function of the rotation of the shaft.

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

If a certain noise is given a loudness of 30 dB, it means it's 30 dB above the threshold of hearing, or 1,000 times as loud as the quietest detectable noise. A noise at 60 dB is 1,000,000 times as powerful as the threshold of hearing. Sound level meters are used to determine the dB levels of various noises and acoustic environments.

A typical conversation might be at a level of about 70 dB. This is 10,000,000 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 times the threshold of hearing.

A sound at 100 dB, typical of the music at a large rock concert, is 10,000,000,000 times as loud, in terms of power, as a sound at the threshold of hearing. If you are sitting in the front row, and if it's a loud band, your ears might get wallopped with peaks of 110 dB. That is 100 billion times the minimum sound power you can detect in a quiet room.

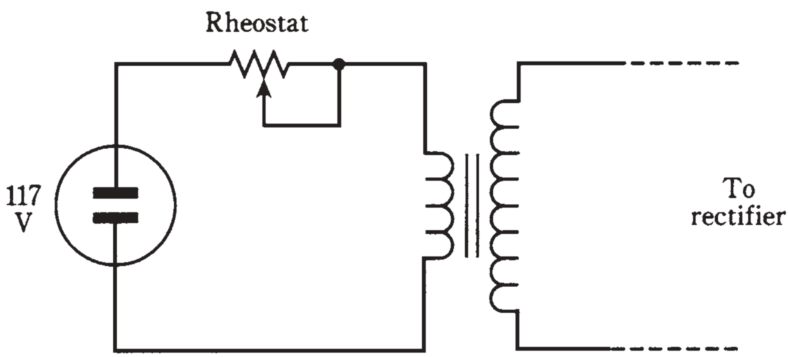
The rheostat

A variable resistor can be made from a wirewound element, rather than a solid strip of material. This is called a *rheostat.* A rheostat can have either a rotary control or a sliding control. This depends on whether the nichrome wire is wound around a doughnut-shaped form *(toroid)* or a cylindrical form (*solenoid*).

Rheostats always have inductance, as well as resistance. They share the advantages and disadvantages of fixed wirewound resistors.

A rheostat is not continuously adjustable, as a potentiometer is. 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. The rheostat resistance therefore adjusts in a series of little jumps.

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. A potentiometer would be destroyed instantly in this application.



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

Resistor values

In theory, a resistor can have any value from the lowest possible (such as a shaft of solid silver) to the highest (open 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 in standard values that might at first seem rather odd to you. The standard numbers are 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. Units are commonly made with values derived from these values, multiplied by some power of 10. Thus, you will see units of 47 Ω, 180 Ω, 6.8 KΩ, or 18 MΩ, but not 380 Ω or 650 KΩ. Maybe you've wondered at some of the resistor values that have been used in problems and quiz questions in previous chapters. Now you know that these choices weren't totally arbitrary; they were picked to represent values you might find in real circuits.

In addition to the above 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.

You don't have to memorize these numbers. They'll become familiar enough over time, as you work with electrical and electronic circuits.

Tolerance

The first set of numbers above represents standard resistance values available in tolerances of plus or minus 10 percent. 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-ohm resistor, for example, the value can be off by as much as 47 ohms and still be within tolerance. That's a range of 423 to 517 ohms. The tolerance is calculated according to the specified value of the resistor, not the actual value. You might measure the value of a 470-ohm resistor and find it to be 427 ohms, and it would be within 10 percent of the specified value; if it measures 420 ohms, it's outside the 10-percent range and is 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. A 470-ohm, 5-percent resistor will have an actual value of 470 ohms plus or minus 24 ohms, or a range of 446 to 494 ohms.

Some resistors are available in tolerances tighter than 5 percent. 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, 10-percent or 5-percent tolerance is good enough. In many cases, even a 20-percent error is all right.

Power rating

All resistors are given a specification that determines how much power they can safely dissipate. Typical values are 1/4 W, 1/2 W, and 1 W. Units also exist with ratings of 1/8 W or 2 W. These dissipation ratings are for continuous duty.

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): $P = I^2R$. Just work this formula backwards, plugging in the power rating for P and the resistance of the unit for R, and solve for I. Or you can find the square root of P/R. Remember to use amperes for current, ohms for resistance, and watts for power.

The power rating for a given resistor can, in effect, be increased by using a network of 2×2, 3×3, 4×4, etc., units in series-parallel. You've already learned about this. If you need a 47-ohm, 45-W resistor, but all you have is a bagful of 47-ohm, 1-W resistors, you can make a 7×7 network in series-parallel, and this will handle 49 W. It might look terrible, but it'll do the job.

Power ratings are specified with a margin for error. A good engineer never tries to take advantage of this and use, say, a 1/4-W unit in a situation where it will need to draw 0.27 W. In fact, good engineers usually include their own safety margin. Allowing 10 percent, a 1/4-W resistor should not be called upon to handle more than about 0.225 W. But it's silly, and needlessly expensive, to use a 2-W resistor where a 1/4-W unit will do, unless, of course, the 2-W resistor is all that's available.

Temperature compensation

All resistors change value somewhat 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 might change. Then the circuit will 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. Usually, it's a needless expense to do this, but if the small change in value cannot be tolerated, it's sometimes the most cost effective.

Still another scheme is to use a series-parallel network of resistors that are all identical, in the manner you already know about, 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 1/4 the intended resistance, and connect them in series.

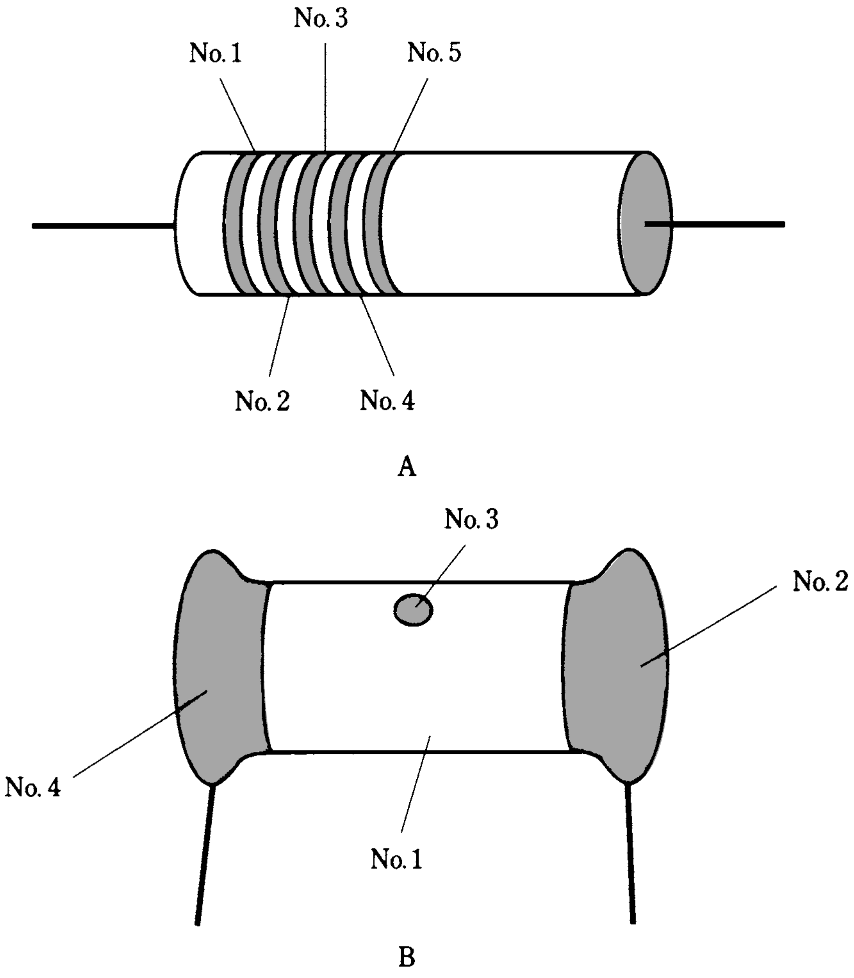
It is unwise to combine several resistors with greatly different values. This can result in one of them taking most of the load while the others loaf, and the combination will be no better than the single hot resistor you started with.

You might get the idea of using two resistors with half (or twice) the value you need, but with *opposite* resistance-versus-temperature characteristics, and connecting them in series (or in parallel). Then the one whose resistance decreases with heat (*negative temperature coefficient)* will have a canceling-out effect on the one whose resistance goes up (*positive temperature coefficient).* This is an elegant theory, but in practice you probably won't be able to find two such resistors without spending at least as much money as you would need to make a 3×3 series-parallel network. And you can't be sure that the opposing effects will exactly balance. It would be better, in such a case, to make a 2×2 series-parallel array of ordinary resistors.

The color code

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,* the bands (first, second, third, fourth, fifth) are arranged as shown in Fig. 6-12A. On resistors with *radial leads,* the bands are arranged



6-12 At A, location of color-code bands on a resistor with axial leads. At B, location of color codings on a resistor having radial leads.

as shown in Fig. 6-12B. The first two bands represent numbers 0 through 9; the third band represents a multiplier of 10 to some power. For the moment, don't worry about the fourth and fifth bands. Refer to Table 6-1.

Table 6-1 Resistor color code

Color of band	Numeral (Bands no.1 and 2.)	Multiplier Band no.3
Black	0	1
Brown	1	10
Red	2	100
Orange	3	1K
Yellow	4	10K
Green	5	100K
Blue	6	1M
Violet	7	10M
Gray	8	100M
White	9	1000M

See text for discussion of bands no. 4 and 5.

Suppose you find a resistor whose first three bands are yellow, violet, and red, in that order. Then the resistance is 4,700 Ω or 4.7 KΩ. Read yellow = 4, violet = 7, red = × 100.

As another example, suppose you stick your hand in a bag and pull out a unit with bands of blue, gray, orange. Refer to Table 6-1 and determine blue = 6, gray = 8, orange = × 1000. Therefore, the value is 68,000 Ω = 68 KΩ.

After a few hundred real-life experiences with this color code, you'll have it memorized. If you aren't going to be using resistors that often, you can always keep a copy of Table 6-1 handy and use it when you need it.

The fourth band, if there is one, indicates tolerance. If it's silver, it means the resistor is rated at plus or minus 10 percent. If it's gold, the resistor is rated at plus or minus 5 percent. If there is no fourth band, the resistor is rated at plus or minus 20 percent.

The fifth band, if there is one, indicates the percentage that the value might change in 1,000 hours of use. A brown band indicates a maximum change of 1 percent of the rated value. A red band indicates 0.1 percent; an orange band indicates 0.01 percent; a yellow band indicates 0.001 percent. If there is no fifth band, it means that the resistor might deviate by more than 1 percent of the rated value after 1,000 hours of use.

A good engineer always tests a resistor with an ohmmeter before installing it. If the unit happens to be labeled wrong, it's easy to catch while assembling a complex electronic circuit. But once the circuit is all together, and it won't work because some resistor is mislabeled (and this happens), it's a gigantic pain to find the problem.

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. Biasing in an amplifier circuit:
   A. Keeps it from oscillating.
   B. Matches it to other amplifier stages in a chain.
   C. Can be done using voltage dividers.
   D. Maximizes current flow.
2. A transistor can be protected from needless overheating by:
   A. Current-limiting resistors.
   B. Bleeder resistors.
   C. Maximizing the driving power.
   D. Shorting out the power supply when the circuit is off.
3. Bleeder resistors:
   A. Are connected across the capacitor in a power supply.
   B. Keep a transistor from drawing too much current.
   C. Prevent an amplifier from being overdriven.
   D. Optimize the efficiency of an amplifier.
4. Carbon-composition resistors:
   A. Can handle lots of power.
   B. Have capacitance or inductance along with resistance.
   C. Are comparatively nonreactive.
   D. Work better for ac than for dc.
5. The best place to use a wirewound resistor is:
   A. In a radio-frequency amplifier.
   B. When the resistor doesn't dissipate much power.
   C. In a high-power, radio-frequency circuit.
   D. In a high-power, direct-current circuit.
6. A metal-film resistor:
   A. Is made using solid carbon/phenolic paste.
   B. Has less reactance than a wirewound type.
   C. Can dissipate large amounts of power.
   D. Has considerable inductance.
7. A meter-sensitivity control in a test instrument would probably be:
   A. A set of switchable, fixed resistors.

B. A linear-taper potentiometer.

C. A logarithmic-taper potentiometer.

D. A wirewound resistor.

8. A volume control in a stereo compact-disc player would probably be:

A. A set of switchable, fixed resistors.

B. A linear-taper potentiometer.

C. A logarithmic-taper potentiometer.

D. A wirewound resistor.

9. If a sound triples in actual power level, approximately what is the decibel increase?

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 W, what is the final sound power?

A. 13 W.

B. 77 mW.

C. 50 mW.

D. There is not enough information to tell.

11. The sound from a transistor 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. 5,000.

D. 100,000.

12. An advantage of a rheostat over a potentiometer is 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:

A. 7.4 percent.

B. 7.9 percent.

C. 5 percent.

D. 10 percent.

14. Suppose a resistor is rated at 3.3 KΩ, plus or minus 5 percent. This means it can be expected to have a value between:

    A. 2,970 and 3,630 Ω.

    B. 3,295 and 3,305 Ω.

    C. 3,135 and 3,465 Ω.

    D. 2.8 KΩ and 3.8 KΩ.

15. A package of resistors is rated at 56 Ω, plus or minus 10 percent. 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 it will have to draw 1 mA maximum continuous current. What power rating is best for this application?

    A. 1/4 W.

    B. 1/2 W.

    C. I W.

    D. 2 W.

17. Suppose a 1-KΩ resistor will dissipate 1.05 W, and you have many 1-W resistors of all common 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. Two 2.2 KΩ, 1-W resistors in parallel.

    C. Three 3.3 KΩ, 1-W resistors in parallel.

    D. One 1 KΩ, 1-W resistor, since manufacturers allow for a 10-percent margin of safety.

18. Red, red, red, gold indicates a resistance of:

    A. 22 Ω.

    B. 220 Ω.

    C. 2.2 KΩ.

    D. 22 KΩ.

19. The actual resistance of the above unit can be expected to vary by how much above or below the specified value?

    A. 11 Ω.

    B. 110 Ω.

    C. 22 Ω.

    D 220 Ω.

20. A resistor has three bands: 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

ONE OF THE MOST COMMON AND MOST VERSATILE SOURCES OF DC IS THE *CELL*. The term *cell* means self-contained compartment, and it can refer to any of various different things in (and out of) science. In electricity and electronics, a cell is a unit source of dc energy. There are dozens of different types of electrical cells.

When two or more cells are connected in series, the result is known as a *battery*.

Kinetic and potential energy

Energy can exist in either of two main forms. *Kinetic energy is* the kind you probably think of right away when you imagine energy. A person running, a car moving down a freeway, a speeding aircraft, a chamber of superheated gas—all these things are visible manifestations of kinetic energy, or energy in action. The dissipation of electrical power, over time, is a form of kinetic energy too.

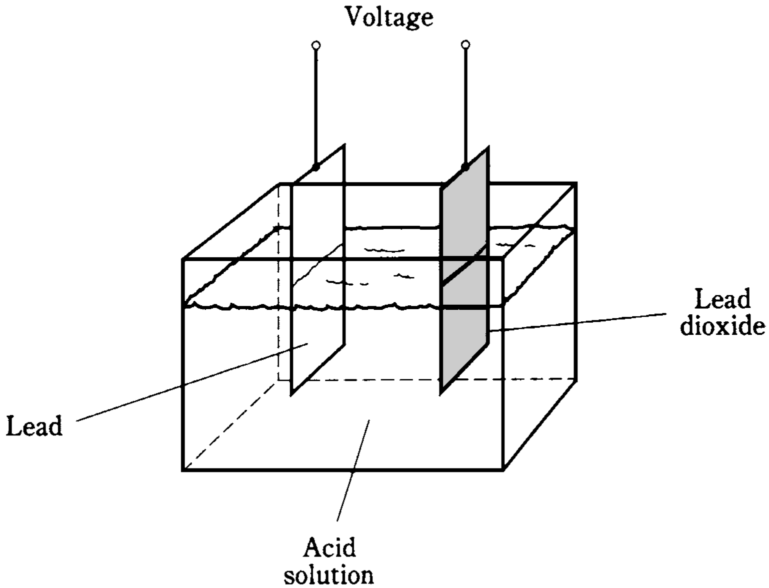
Potential energy is not as vividly apparent. When you raise a block of concrete into the air, you are creating potential energy. You remember the units called *foot pounds,* the best way to measure such energy, from school physics classes. If you raise a one-pound weight a foot, it gains one foot pound of potential energy. If you raise it 100 feet, it gains 100 foot pounds. If you raise a 100-pound weight 100 feet, it will gain 100 × 100, or 10,000, foot pounds of potential energy. This energy becomes spectacularly evident if you happen to drop a 100-pound weight from a tenth-story window. (But don't!)

Electrochemical energy

In electricity, one important form of potential energy exists in the atoms and molecules of some chemicals under special conditions.

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) will show a persistent voltage. This can be detected by connecting a galvanometer between the pieces of metal. A resistor of about 1,000 ohms should *always* be used in series with the galvanometer in experiments of this kind; connecting the galvanometer directly will cause too much current to flow, possibly damaging the galvanometer and causing the acid to "boil."



7-1 Construction of a lead-acid electrochemical cell.

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. The acid will change, also. The *chemical energy,* a form of potential energy in the acid, will run out. All of the potential energy in the acid will have been turned into kinetic 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 potential (chemical) energy has all been changed to electricity and used up, must be thrown away. They are no good anymore. These are called *primary cells.*

Other kinds of cells, like the lead-and-acid unit depicted above, can get their chemical energy back again. 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, alkaline cell,* and others. Go into a department store and find the panel of batteries, and you'll see various sizes and types of primary cells, such as AAA batteries, D batteries, camera batteries, and watch batteries. You should know by now that these things are cells, not true batteries. This is a good example of a misnomer that has gotten so widespread that store clerks might look at you funny if you ask for a couple of cells. You'll also see real batteries, such as the little 9-V transistor batteries and the large 6-V lantern batteries.

Secondary cells can also be found increasingly in consumer stores. *Nickel-cadmium (Ni-Cd or NICAD)* cells are probably the most common. They're available in some of the same sizes as nonrechargeable dry cells. 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 your 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 extremely dangerous to short-circuit the terminals of such a battery, because the acid (sulfuric acid) can "boil" out and burn your skin and eyes.

An important note is worth making here: *Never* short-circuit any cell or battery, because it might burst or explode.

The Weston standard cell

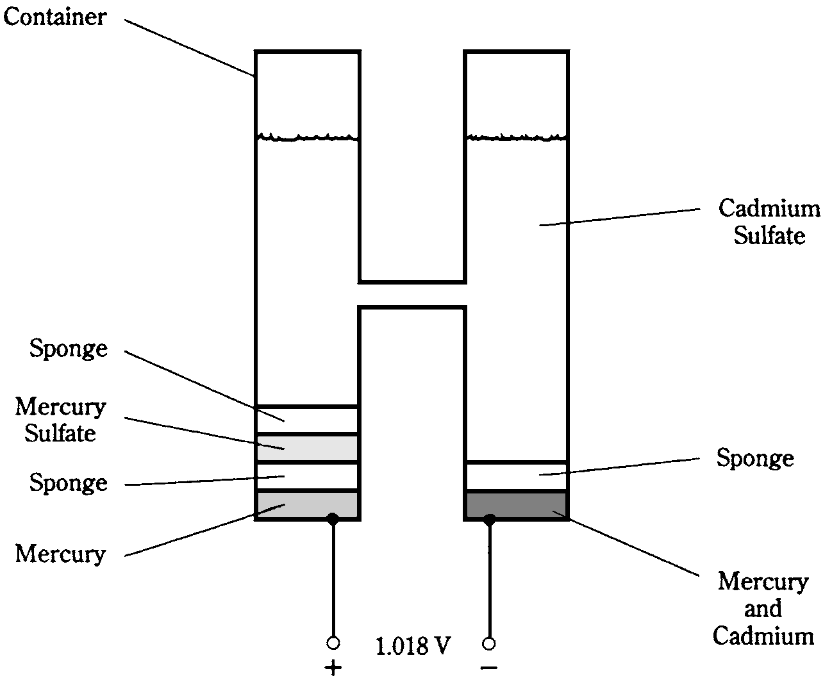
Most electrochemical cells produce about 1.2 V to 1.8 V of electric potential. Different types vary slightly. A mercury cell has a voltage that is a little bit less than that of a zinc-carbon or alkaline cell. The voltage of a cell can also be affected by variables in the manufacturing process. Normally, this is not significant. Most consumer type dry cells can be assumed to produce 1.5 Vdc.

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

When properly constructed and used at room temperature, the voltage of the Weston standard cell is always the same, and this allows it to be used as a dc voltage standard. There are other kinds of standard cells, but the Weston cell is the most common.

Storage capacity

Recall that the unit of energy is the watt hour (Wh) or the kilowatt hour (kWh). Any electrochemical cell or battery has a certain amount of electrical energy that can be gotten



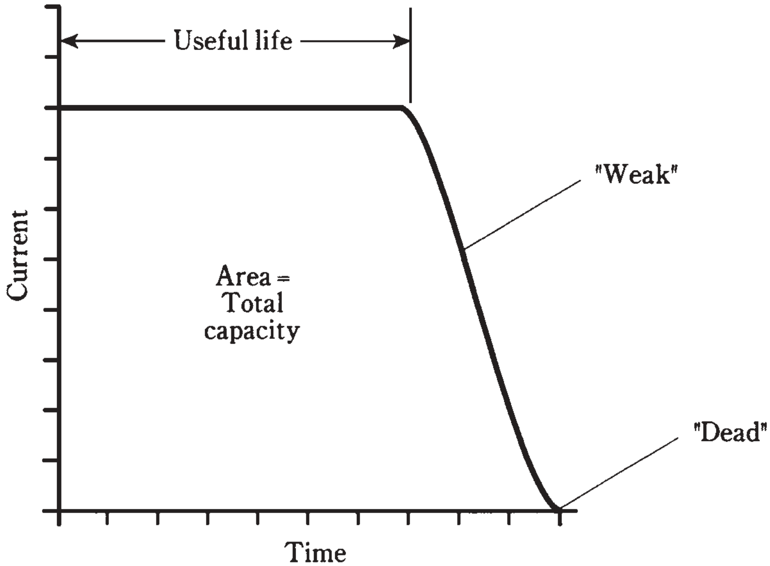
7-2 A Weston standard cell.

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 an hour, or 1 A for 2 hours. Or it can provide 100 mA for 20 hours. Within reason, the product of the current in amperes, and the use time in hours, can be as much as, but not more than 2. The limitations are the *shelf life* at one extreme, and the *maximum safe current* at the other. Shelf life is the length of time the battery will last if it is sitting on a shelf without being used; this might be years. The maximum safe current is represented by the lowest load resistance (heaviest load) that the battery can work into before its voltage drops because of its own *internal resistance.* A battery is never used with loads that are too heavy, because it can't supply the necessary current anyway, and it might "boil", burst, or blow up.

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

When an ideal cell or battery is used, it delivers a fairly constant current for awhile, and then the current starts to fall off (Fig. 7-3). Some types of cells and batteries approach this ideal behavior, called a *flat discharge curve,* and others have 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." Some rechargeable cells and batteries, especially the nickel-cadmium type, should never be used until the current goes down to zero, because this can ruin them.



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

The area under the curve in Fig. 7-3 is the total capacity of the cell or battery in ampere hours. This area is always pretty much the same for any particular type and size of cell or battery, regardless of the amount of current drawn while it's in use.

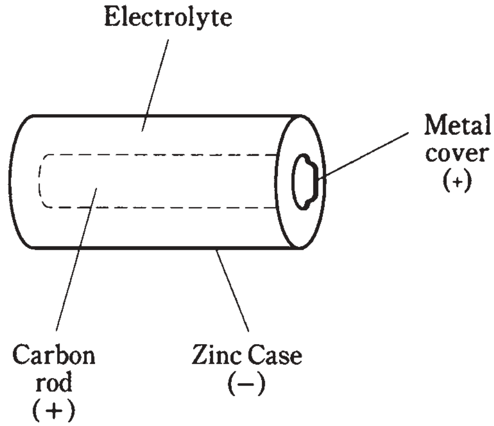
Common dime-store cells and batteries

The cells you see in grocery, department, drug, and hardware stores that are popular for use in household convenience items like flashlights and transistor radios are usually of the *zinc-carbon* or *alkaline* variety. These provide 1.5 V and are available in sizes known as AAA (very small), AA (small), C (medium large), and D (large). You have probably seen all of these sizes hanging in packages on a pegboard. Batteries made from these cells are usually 6 V or 9 V.

One type of cell and battery that has become available recently, the nickel-cadmium rechargeable type, is discussed in some detail a bit later in this chapter.

Zinc-carbon cells

These cells have a fairly long shelf life. A cylindrical zinc-carbon cutaway diagram is shown at Fig. 7-4. 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.



7-4 Simplified diagram of zinc-carbon cylindrical cell construction.

Alkaline cells

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

Those little 9-V things with the funny connectors on top consist of six tiny zinc-carbon or alkaline cells in series. Each of the six cells supplies 1.5 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 gotten 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* (size) of the cell. C or D size cells have more volume than a transistor battery, and therefore contain more stored energy, assuming the same chemical type.

The ampere-hour capacity of a transistor battery is very small. But transistor radios don't need much current. These batteries are also used in other low-current electronic devices, such as remote-control garage-door openers, TV channel changers, remote video-cassette recorder (VCR) controls, and electronic calculators.

Lantern batteries

These get their name from the fact that they find much of their use in lanterns. These are the batteries with a good, solid mass so they last a long time. One type has spring contacts on the top. The other type has thumbscrew terminals. Besides keeping a lantern lit for awhile, these big batteries, usually rated at 6 V and consisting of four good-size zinc-carbon or alkaline cells, can provide enough energy to operate a low-power radio transceiver. Two of them in series make a 12-V battery that can power a 5-W Citizen

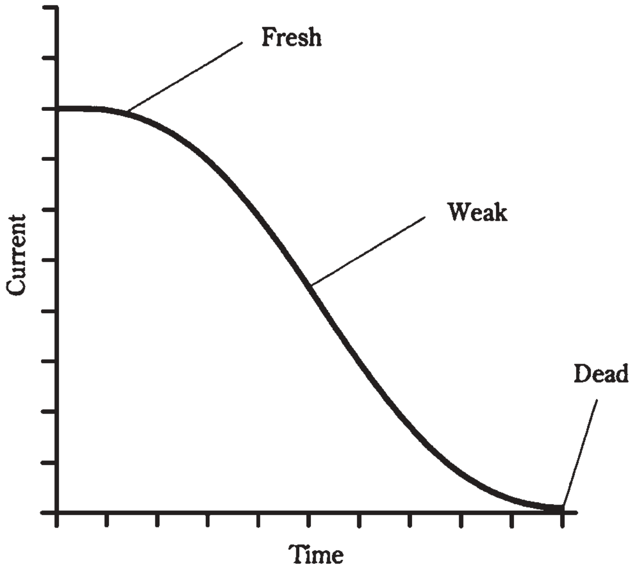
Band (CB) or ham radio. They're also good for scanner radio receivers in portable locations, for camping lamps, and for other medium-power needs.

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 watches, cameras, and other microminiature electronic gizmos.

Silver-oxide types

Silver-oxide cells are usually made into button-like shapes, and can fit inside even a small wristwatch. They 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 flat discharge curve, like the one shown in the graph of Fig. 7-3. The previously described zinc-carbon and alkaline cells and batteries have a current output that declines with time in a steady fashion, as shown in Fig. 7-5. This is known as a *declining discharge curve.*



7-5 A declining discharge curve.

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

Mercury types

Mercury cells, also called *mercuric oxide* cells, have advantages similar to 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 resulting 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 some decrease in the popularity of mercury cells and batteries in recent years. This is because of the fact that mercury is highly toxic. When mercury cells and batteries are dead, they must be discarded. Eventually the mercury, a chemical element, leaks into the soil and ground water. Mercury pollution has become a significant concern in places that might really surprise you.

Lithium types

Lithium cells have become popular since the early eighties. 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 V to 3.5 V, depending on the particular chemistry used. These cells, like their silver-oxide cousins, can be stacked to make batteries.

The first applications 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 energy capacity per unit volume that is vastly greater than other types of electrochemical cells.

Lithium cells and batteries are used in low-power devices that must operate for a long time without power-source replacement. Heart pacemakers and security systems are two examples of such applications.

Lead-acid cells and batteries

You've already 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 batteries 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 get out.

Lead-acid batteries are also available in a construction that uses a semisolid electrolyte. These batteries are popular in consumer electronic devices that require a moderate amount of current. Notebook or laptop computers, and portable video-cassette recorders (VCRs), are the best examples.

A large lead-acid battery, such as the kind in your car, can store several tens of ampere-hours. The smaller ones, like those in notebook computers, have less capacity but more versatility. Their overwhelming advantage is their ability to be used many times at reasonable cost.

Nickel-cadmium cells and batteries

You've probably seen, or at least heard of, *NICAD* cells and batteries. They have become quite common in consumer devices such as those little radios and cassette players you can wear while doing aerobics or just sitting around. (These entertainment units are not too safe for walking or jogging in traffic. And never wear them while riding a bicycle.) You can buy two sets of cells and switch them every couple of hours of use, charging one set while using the other. Plug-in charger units cost only a few dollars.

Types of NICAD cells

Nickel-cadmium cells are made in several types. *Cylindrical cells* are the standard cells; they look like dry cells. *Button cells* are those little things that are used in cameras, watches, memory backup applications, and other places where miniaturization is important. *Flooded cells* are used in heavy-duty applications and can have a charge capacity of as much as 1,000 Ah. *Spacecraft cells* are made in packages that can withstand the vacuum and temperature changes of a spaceborne environment.

Uses of NICADs

There are other uses for NICADs besides in portable entertainment equipment. 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 set of NICADs, in addition to powering the satellite, for half of each orbit. The NICADs can provide the power during the dark half of each orbit.

Nickel-cadmium batteries are available in packs of cells. These packs can be plugged into the equipment, and might even form part of the case for a device. An example of this is the battery pack for a handheld ham radio tranceiver. Two of these packs can be bought, and they can be used alternately, with one installed in the "handie-talkie" (HT) while the other is being charged.

NICAD neuroses

There are some things you need to know about NICAD cells and batteries, in order to get the most out of them.

One rule, already mentioned, is that you should never discharge them all the way until they "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.

Another phenomenon, peculiar to this type of cell and battery, is called *memory.* If a NICAD is used over and over, and is discharged to exactly the same extent every time (say, two-thirds of the way), it might start to "go to sleep" at that point in its discharge cycle. This is uncommon; lab scientists have trouble forcing it to occur so they can study it. But when it does happen, it can give the illusion that the cell or battery has lost some of its storage capacity. Memory problems can be solved. Use the cell or battery almost all the way up, and then fully charge it. Repeat the process, and the memory will be "erased."

NICADS do best using wall chargers that take several hours to fully replenish the cells or batteries. There are *high-rate* or *quick* chargers available, but these can sometimes force too much current through a NICAD. 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 Radio Shack store, should be able to tell you which chargers are best for which cells and batteries.

Photovoltaic cells and batteries

The *photovoltaic* cell is completely different from any of the electrochemical cells. It's also known as a *solar* cell. This device converts visible light, infrared, and/or ultraviolet 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

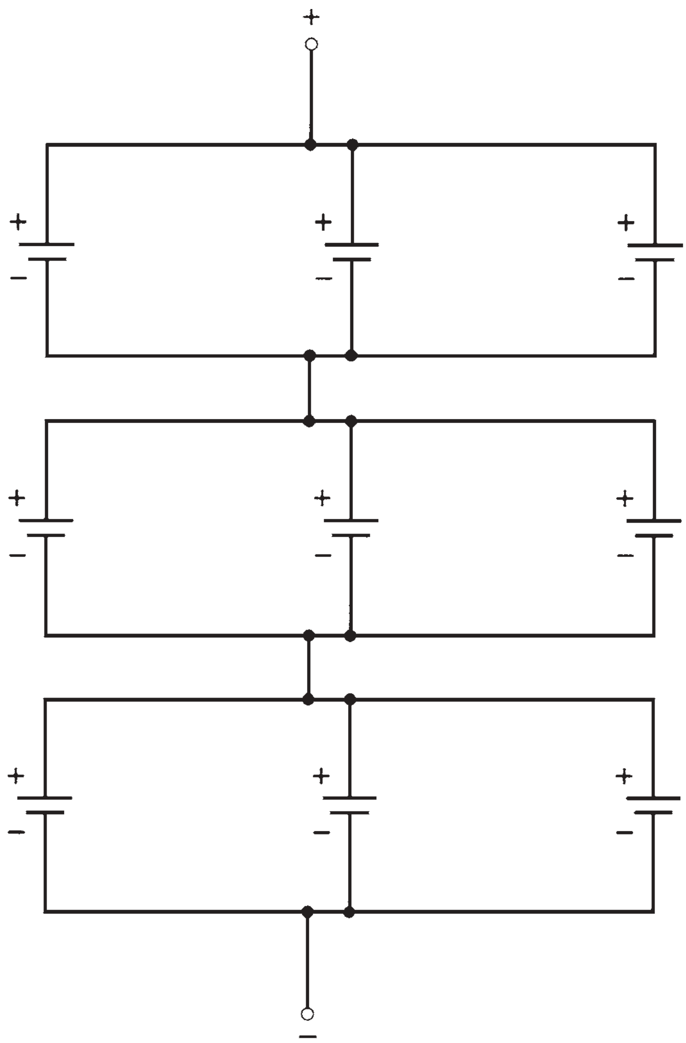
The construction of a photovoltaic cell is shown in Fig. 7-7. The device is a flat semiconductor *P-N junction,* and the assembly is made transparent so that light can fall directly on the *P-type* silicon. The metal ribbing, forming the positive electrode, is interconnected by means of tiny wires. The negative electrode is a metal backing, 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 a long-term 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 generally uses large solar panels, large-capacity lead-acid or NICAD batteries, power converters to convert the dc into ac, and a rather sophisticated charging circuit. These systems are best suited to environments where there is sunshine a high percentage of the time.

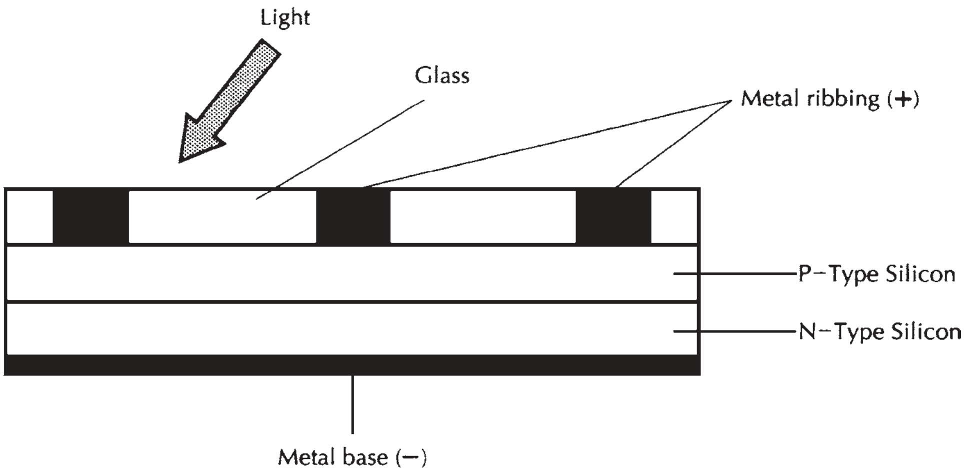
Solar cells, either alone or supplemented with rechargeable batteries, can be connected into a home electric system in an *interactive* arrangement with the electric utilities. When the solar power system can't provide for the needs of the household all by itself, the utility company can take up the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



7-6 Connection of cells in series-parallel.

How large of a battery?

You might get the idea that you can connect hundreds, or even thousands, of cells in series and obtain batteries with fantastically high EMFs. Why not put 1,000 zinc-carbon cells in series, for example, and get 1.5 kV? Or put 2,500 solar cells in series and get 1.25 kV? Maybe it's possible to put a billion solar cells in series, out in some vast sun-scorched desert wasteland, and use the resulting 500 MV (megavolts) to feed the greatest high-tension power line the world has ever seen.



7-7 Cross-sectional view of silicon photovoltaic (solar) cell construction.

There are several reasons why these schemes aren't good ideas. First, high voltages for practical purposes can be generated cheaply and efficiently by power converters that work from 117-V or 234-V utility mains. Second, it would be difficult to maintain a battery of thousands, millions or billions of cells in series. Imagine a cell holder with 1,000 sets of contacts. And not one of them can open up, lest the whole battery become useless, because all the cells must be in series. (Solar panels, at least, can be permanently wired together. Not so with batteries that must often be replaced.) And finally, the *internal resistances* of the cells would add up and limit the current, as well as the output voltage, that could be derived by connecting so many cells in series. This is not so much of a problem with series-parallel combinations, as in solar panels, as long as the voltages are reasonable. But it is a big factor if all the cells are in series, with the intent of getting a huge voltage. This effect will occur with any kind of cell, whether electrochemical or photovoltaic.

In the days of the Second World War, portable two-way radios were built using vacuum tubes. These were powered by batteries supplying 103.5 V. The batteries were several inches long and about an inch in diameter. They were made by stacking many little zinc-carbon cells on top of each other, and enclosing the whole assembly in a single case. You could get a nasty jolt from one of those things. They were downright dangerous! A fresh 103.5-V battery would light up a 15-W household incandescent bulb to almost full brilliance. But the 117-V outlet would work better, and for a lot longer.

Nowadays, handheld radio transceivers will work from NICAD battery packs or batteries of ordinary dry cells, providing 6 V, 9 V, or 12 V. Even the biggest power transistors rarely use higher voltages. Automotive or truck batteries can produce more than enough power for almost any mobile or portable communications system. And if a really substantial setup is desired, gasoline-powered generators are available, and they will supply the needed energy at far less cost than batteries. There's just no use for a megabattery of a thousand, a million, or a zillion volts.

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 kinetic energy when the cell is used.
   D. Is caused by electric current.

2. A cell that cannot be recharged is:
   A. A dry cell.
   B. A wet cell.
   C. A primary cell.
   D. A 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 an energy reference source.

4. The voltage in a battery is:
   A. Less than the voltage in a cell of the same kind.
   B. The same as the voltage in a cell of the same kind.
   C. More than the voltage in a cell of the same kind.
   D. Always a multiple of 1.018 V.

5. A direct short-circuit of a 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. An explosion.

6. A cell of 1.5 V supplies 100 mA for seven hours and twenty minutes, and then it is replaced. It has supplied:
   A. 7.33 Ah.
   B. 733 mAh.
   C. 7.33 Wh.
   D. 733 mWh.

7. A 12-V auto battery is rated at 36 Ah. If a 100-W, 12-Vdc bulb is connected across this battery, about how long will the bulb stay lit, if 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. Are generally 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 places 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 places would a transistor battery be the best power-source choice?
    A. A heart pacemaker.
    B. An electronic calculator.
    C. An LCD wristwatch.
    D. A two-way portable radio.

12. In which of the following places would you most likely choose a lithium battery?
    A. A microcomputer memory backup.
    B. A two-way portable radio.
    C. A portable audio cassette player.
    D. A rechargeable flashlight.

13. Where would you most likely find a lead-acid battery?
    A. In a portable audio cassette player.

B. In a portable video camera/recorder.

C. In an LCD wall clock.

D. In a flashlight.

14. A cell or battery that keeps up 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 per unit volume.

15. Where might you find a NICAD battery?

A. In a satellite.

B. In a portable cassette player.

C. In a handheld radio transceiver.

D. In more than one of the above.

16. A disadvantage of mercury cells and batteries is that:

A. They don't last as long as other types.

B. They have a flat discharge curve.

C. They pollute 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-cadmium.

D. Mercury.

18. The current from a solar panel is increased by:

A. Connecting solar cells in series.

B. Using NICAD cells in series with the solar cells.

C. Connecting solar cells in parallel.

D. Using lead-acid cells in series with the solar cells.

19. An interactive solar power system:

A. Allows a homeowner to sell power to the utility.

B. Lets the batteries recharge at night.

C. Powers lights but not electronic devices.

D. Is totally independent from the utility.

20. One reason why it is impractical to make an extrememly high-voltage battery of cells is that:

    A. There's a danger of electric shock.

    B. It is impossible to get more than 103.5 V with electrochemical cells.

    C. The battery would weigh to much.

    D. There isn't any real need for such thing.

8
CHAPTER
Magnetism

THE STUDY OF MAGNETISM IS A SCIENCE IN ITSELF. ELECTRIC AND MAGNETIC phenomena interact; a detailed study of magnetism and electromagnetism could easily fill a book. Magnetism was mentioned briefly near the end of chapter 2. Here, the subject is examined more closely. The intent is to get you familiar with the general concepts of magnetism, insofar as it is important for a basic understanding of electricity and electronics.

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 huge magnetic field that surrounds the earth. The sun has a magnetic field, as does the whole Milky Way galaxy. These fields might have originally magnetized the earth.

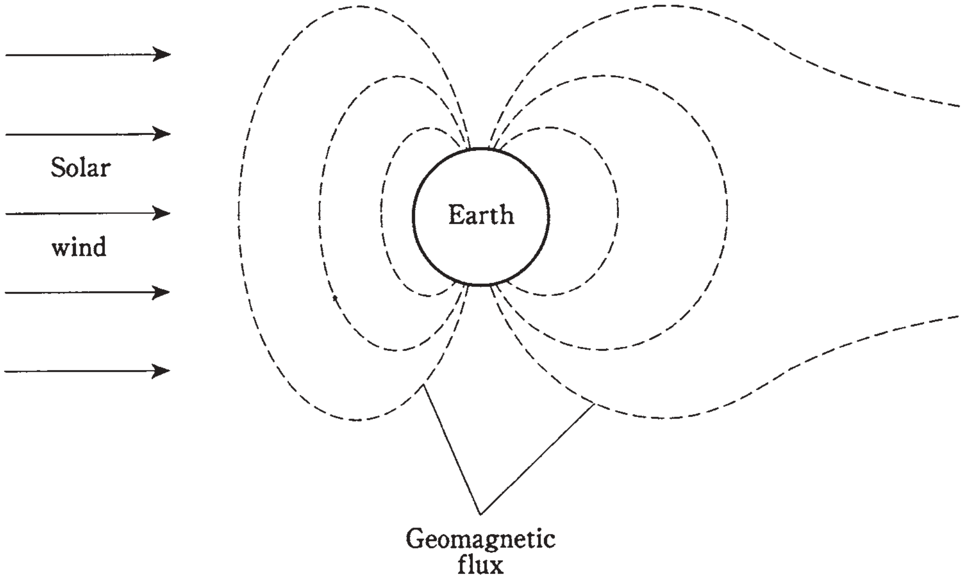
Geomagnetic poles and axis

The *geomagnetic field,* as it is called, has poles, just as a bar magnet does. These poles are near, but not at, the geographic poles. The north geomagnetic pole is located in the frozen island region of northern Canada. The south geomagnetic pole is near Antarctica. The *geomagnetic axis* is somewhat tilted relative to the axis on which the earth rotates. Not only this, but it does not exactly run through the center of the earth. It's like an apple core that's off center.

The solar wind

The geomagnetic field would be symmetrical around the earth, but charged particles from the sun, constantly streaming outward through the solar system, distort the lines of flux.

This *solar wind* literally "blows" the geomagnetic field out of shape, as shown in Fig. 8-1. At and near the earth's surface, the lines of flux are not affected very much, and the geomagnetic field is nearly symmetrical.



8-1 The geomagnetic field is distorted by the solar wind.

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 attributed to the presence of a "force" in the air. Even though it was some time before the details were fully understood, 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 and the magnetic field around a compass needle interact, so that a force is exerted on the little magnet inside the compass. 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 lattitude increases, either towards the north or the south geomagnetic pole, the magnetic force pulls up and down on the compass needle more and more. You have probably noticed this when you hold a compass. 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 magnetic *lines of flux.*

Magnetic force

Magnets "stick" to some metals. Iron, nickel, and alloys containing either or both of these elements, are known as *ferromagnetic* materials. When a magnet is brought near a piece of ferromagnetic material, the atoms in the material become lined up, so that the metal is temporarily magnetized. This produces a *magnetic force* between the atoms of the ferromagnetic substance and those in the magnet.

If a magnet is brought near another magnet, the force is even stronger. Not only is it more powerful, but it can be repulsive or attractive, depending on the way the magnets are turned. The force gets stronger as the magnets are brought near each other.

Some magnets are so strong that no human being can ever 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 available are of use in industry. A huge 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 is called *magnetic levitation* and is the basis for some low-friction, high-speed trains now being developed.

Electric charge in motion

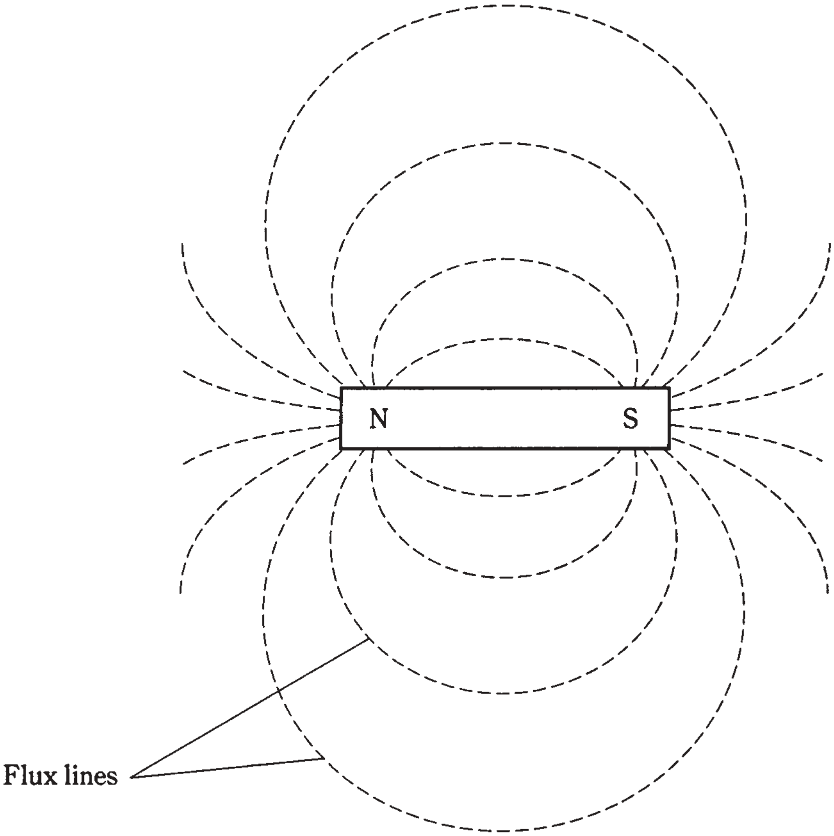
Whenever the atoms in a ferromagnetic material are aligned, a *magnetic field* exists. A magnetic field can also be caused by the motion of electric charge carriers, either in a wire or in free space.

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, the 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 a sort of current is produced just by the way they move within individual atoms.

Magnetic fields can be produced 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, they have magnetic fields. When these fields interact with the geomagnetic field, the particles are forced to change direction. Charged particles from the sun are *accelerated* toward the geomagnetic poles. If there is a *solar flare,* the sun ejects far more charged particles than normal. When these arrive at the geomagnetic poles, the result can actually disrupt the geomagnetic field. Then there is a *geomagnetic storm.* This 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. Microwave transmissions are generally immune to the effects of a geomagnetic storm, although the wire links can be affected. Aurora (northern or southern lights) are frequently observed at night during these events.

Flux lines

Perhaps you have seen the experiment in which iron filings are placed on a 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).



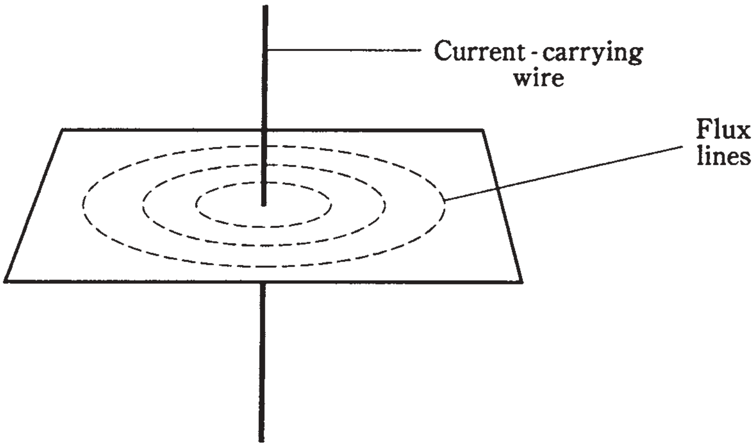
8-2 Pattern of magnetic flux lines around a bar magnet.

Another experiment involves passing a current-carrying wire through the paper at a right angle, as shown in Fig. 8-3. The iron filings will be grouped along circles centered at the point where the wire passes through the paper.

Physicists consider magnetic fields to be comprised of flux lines. The intensity of the 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 really exist as geometric threads in space, or as anything solid, but it is intuitively appealing to imagine them, and the iron filings on the paper really do bunch themselves into lines when there is a magnetic field of sufficient strength to make them move.

Magnetic polarity

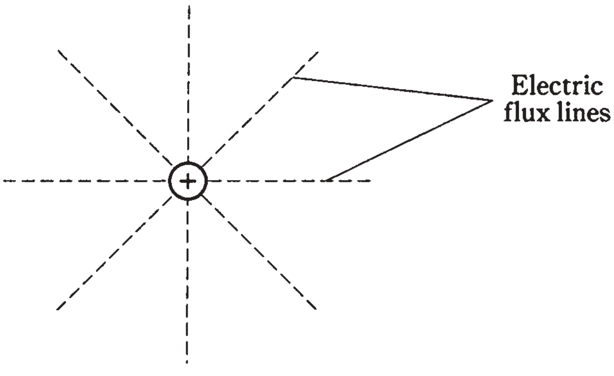
A magnetic field has a direction at any given point in space near a current-carrying wire or a permanent magnet. The flux lines run parallel with the direction of the field. A magnetic field is considered to begin at the north magnetic pole, and to terminate at the south magnetic pole. In the case of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



8-3 Pattern of magnetic flux lines around a current-carrying wire.

With a current-carrying wire, the magnetic field just goes around and around endlessly, like a dog chasing its own tail.

A charged electric particle, such as a proton, hovering in space, is a *monopole,* and the electric flux lines around it aren't closed (Fig. 8-4). A positive charge does not have to be mated with a negative charge. The electric flux lines around any stationary, charged particle run outward in all directions for a theoretically infinite distance.



8-4 Electric flux lines around a monopole charge.

But a magnetic field is different. All magnetic flux lines are *closed loops.* With permanent magnets, there is always a starting point (the north pole) and an ending point (the south pole). Around the current-carrying wire, the loops are circles. This can be plainly seen in the experiments with iron filings on paper. Never do magnetic flux lines run off into infinity. Never is a magnetic pole found without an accompanying, opposite pole.

Dipoles and monopoles

A pair of magnetic poles is called a *dipole.* A lone pole, like the positive pole in a proton, is called a *monopole.*

Magnetic monopoles do not ordinarily exist in nature. If they could somehow be conjured up, all sorts of fascinating things might happen. Scientists are researching this to see if they can create artificial magnetic monopoles.

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 in fact, you can think of any half-plane, with the edge along the line of the wire, as a magnetic dipole, and the lines of flux as going around once from the "north face" of the half-plane to the "south face."

The lines of flux in a magnetic field always connect the two poles. Some flux lines are straight; some are curved. 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.

Magnetic field strength

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

The tesla and the gauss

If you have a certain permanent magnet or electromagnet, you might see its "strength" expressed in terms of webers or maxwells. But usually you'll hear units called teslas or gauss. 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, are more useful expressions for magnetic effects than the overall quantity of magnetism. A flux density of one tesla is equal to one weber per square meter. A flux density of one gauss is equal to one maxwell per square centimeter. It turns out that the gauss is 0.0001, or 10^{-4}, tesla. Conversely, the tesla is 10,000, or 10^4, gauss.

If you are confused by the distinctions between webers and teslas, or between maxwells and gauss, think of a light bulb. A 100-watt lamp might emit a total of 20 watts of visible-light power. If you enclose the bulb completely, then 20 W will fall on the interior walls of the chamber, no matter how large or small the chamber might be. But this is not a very useful notion of the brightness of the light. You know that a single 100-watt light bulb gives plenty of light for a small walk-in closet, but it is nowhere near adequate to illuminate a gymnasium. The important consideration is the number of watts per unit area. When we say the bulb gives off x watts or milliwatts of light, it's like saying a magnet has y webers or maxwells of magnetism. When we say that the bulb provides x watts or milliwatts per square meter, it's analogous to saying that a magnetic field has a flux density of y teslas or gauss.

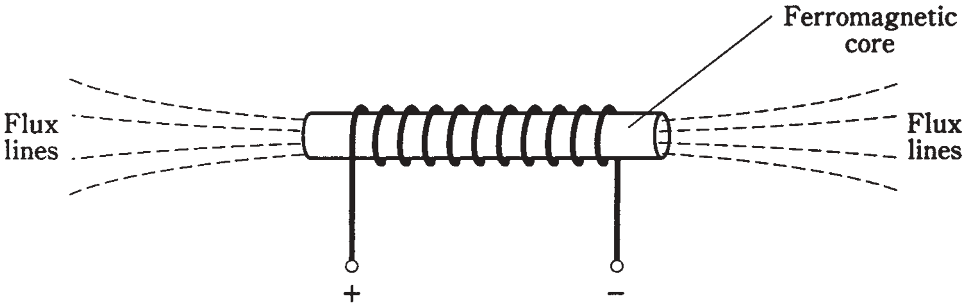
The ampere-turn and the gilbert

When working with electromagnets, another unit is employed. This is the *ampere-turn* (At). It is a unit of *magnetomotive force.* A wire, bent into a circle and carrying 1 A of current, will produce 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 will be 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* is also sometimes used to express magnetomotive force. This unit is equal to 0.796 At. Thus, to get ampere-turns from gilberts, multiply by 0.796; to get gilberts from ampere-turns, multiply by 1.26.

Electromagnets

Any electric current, or movement of charge carriers, produces a magnetic field. This field can become extremely intense in a tightly coiled wire having many turns, and that carries a large electric 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 becomes tremendous. This is the principle of an *electromagnet* (Fig. 8-5).



8-5 A ferromagnetic core concentrates the lines of magnetic flux.

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 always the same: the magnetic field produced by the current results in magnetization of the core.

Direct-current types

Electromagnets can use either direct or alternating current. The type with which you are probably familiar is the dc electromagnet.

You can build a dc electromagnet by taking a large bolt, like a stove bolt, and wrapping a few dozen or a few hundred turns of wire around it. These items are available in almost any 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 3/8 inch in diameter and several inches long. You need to use insulated wire, preferably made of solid, soft copper. "Bell wire" works very well.

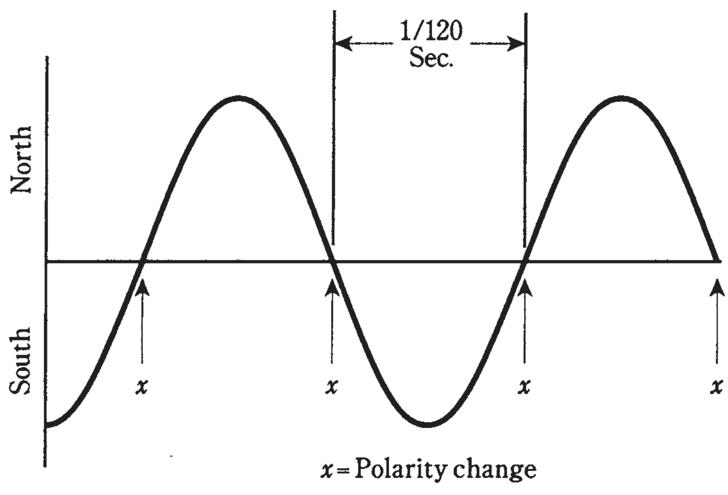
Just be sure all the wire turns go in the same direction. A large 6-V lantern battery can provide plenty of dc to work the electromagnet. Never leave the coil connected to the battery for more than a few seconds at a time. And never use a car battery for this experiment! (The acid might boil out.)

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 should see evidence of this if you do the above experiment with a large enough bolt and enough turns of wire.

Aternating-current types

You might get the idea that the 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 practice, you'll probably 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 117-V 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. That means there are 120 fluctuations per second, or 60 complete north-to-south-to-north polarity changes (Fig. 8-6) every second. If a permanent magnet, or a dc electromagnet, is brought near either "pole" of an ac electromagnet, there will be no net force. This is because the poles will be alike half the time, and opposite half the time, producing an equal amount of attractive and repulsive force.



8-6 Polarity change in an ac electromagnet.

For an ac electromagnet to work, the core material must have high *permeability* but *low retentivity.* These terms will now be discussed.

Permeability

Some substances cause the magnetic lines of flux to get closer together than they are in the air. Some materials can cause the lines of flux to become farther apart than they are in the air.

The first kind of material is *ferromagnetic,* and is of primary importance in magnetism. Ferromagnetic substances are the ones that can be "magnetized." Iron and nickel are examples. Various alloys are even more ferromagnetic than pure iron or pure nickel.

The other kind of material is called *diamagnetic.* Wax, dry wood, bismuth, and silver are substances that actually decrease the magnetic flux density. No diamagnetic material reduces the strength of a magnetic field by anywhere near the factor that ferromagnetic substances can increase it.

Permeability 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 factor of about 60 to several thousand times. Therefore, the permeability of iron can range from 60 (impure) to as much as 8,000 (highly refined).

If you use certain *permalloys* as the core material in electromagnets, you can increase the flux density, and therefore the local strength of the field, by as much as 1,000,000 times. Such substances thus have permeability as great as 1,000,000.

If for some reason you feel compelled to make an electromagnet that is as weak as possible, you could use dry wood or wax for the core material. But usually, diamagnetic substances are used to keep magnetic objects apart, while minimizing the interaction between them.

Diamagnetic metals have the useful property that they conduct electric current very well, but magnetic current very poorly. They can be used for *electrostatic shielding,* a means of allowing magnetic fields to pass through while blocking electric fields.

Table 8-1 gives the permeability ratings for some common materials.

Retentivity

Certain ferromagnetic materials stay magnetized better than others. 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 when the current stops flowing in the coil. *Retentivity,* also sometimes called *remanence,* is a measure of how well the substance will "memorize" the magnetism, and thereby become a permanent magnet.

Retentivity is expressed as a percentage. 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 $100(y/x)$.

As an example, suppose that a metal rod can be magnetized to 135 gauss 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

Table 8-1. Permeability of some common materials.

Substance	Permeability (approx.)
Aluminum	Slightly more than 1
Bismuth	Slightly less than 1
Cobalt	60-70
Ferrite	100-3000
Free space	1
Iron	60-100
Iron, refined	3000-8000
Nickel	50-60
Permalloy	3000-30,000
Silver	Slightly less than 1
Steel	300-600
Super permalloys	100,000-1,000,000
Wax	Slightly less than 1
Wood, dry	Slightly less than 1

maximum; further increasing the current in the wire will not make the rod any more magnetic. Now suppose that the current is shut off, and 19 gauss remain in the rod. Then the retentivity, B_r, is

$$B_r = 100(19/135) = 100 \times 0.14 = 14 \text{ percent}$$

Various different substances have good retentivity; these are excellent for making permanent magnets. Other materials have poor retentivity. They might work well as electromagnets, but not as permanent magnets.

Sometimes it is desirable to have a substance with good ferromagnetic properties, but poor retentivity. This is the case when you want to have an electromagnet that will operate from dc, so that it maintains a constant polarity, but that will lose its magnetism when the current is shut off.

If a ferromagnetic substance has poor retentivity, it's easy to make it work as the core for an ac electromagnet, because the polarity is easy to switch. If the retentivity is very high, the material is "sluggish" and will not work well for ac electromagnets.

Permanent magnets

Any ferromagnetic material, or substance whose atoms can be permanently aligned, can be made into a permanent magnet. These are the magnets you probably played with as a child. Some alloys can be made into stronger magnets than others.

One alloy that is especially suited to making strong magnets is *alnico*. This word derives from the metals that comprise it: *al*uminum, *ni*ckel and *co*balt. Other elements are often added, including copper and titanium. But any piece of iron or steel can be magnetized, at least to some extent. You might have used a screwdriver, for example, that was magnetized, so that it could hold on to screws when installing or removing them from hard-to-reach places.

Permanent magnets are best made from materials with high retentivity. Magnets are made by using a high-retentivity ferromagnetic material as the core of an electromagnet for an extended period of time. This experiment is not a good one to do at home with a battery, because there is a risk of battery explosion.

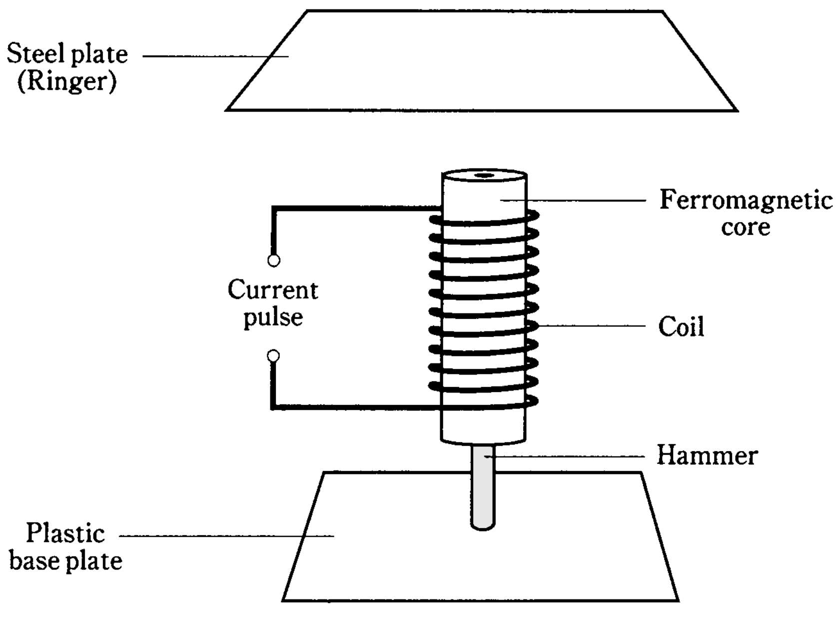
If you want to magnetize a screwdriver a little bit so that it will hold onto screws, just stroke the shaft of the screwdriver with the end of a bar magnet several dozen times. But remember that once you have magnetized a tool, it is difficult to completely demagnetize it.

The solenoid

A cylindrical coil, having a movable ferromagnetic core, can be useful for various things. This is a *solenoid.* Electrical relays, bell ringers, electric "hammers," and other mechanical devices make use of the principle of the solenoid.

A ringer device

Figure 8-7 is a simplified diagram of a bell ringer. Its solenoid is an electromagnet, except that the core is not completely solid, but 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 quite powerful. A movable steel rod runs through the hole in the electromagnet core.



8-7 A solenoid-coil bell ringer.

When there is no current flowing in the coil, the steel rod is held down by the force of gravity. But 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 that the steel rod falls back down again to its resting position, allowing the plate to reverberate: *Gonggg!* Some office telephones are equipped with ringers that produce this noise, rather than the conventional ringing or electronic bleeping ernitted by most phone sets.

A relay

In some electronic devices, it is inconvenient to place a switch exactly where it should be. For example, you might want to switch a communications line from one branch to another from a long distance away. In many radio transmitters, the wiring carries high-frequency alternating currents that must be kept within certain parts of the circuit, and not routed out to the front panel for switching. A *relay* makes use of a solenoid to allow remote-control switching.

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 used for different purposes. Some are meant for use with dc, and others are for ac; a few will work with either type of current. A *normally closed* relay completes the circuit when there is no current flowing in its electromagnet, and breaks the circuit when current flows. A *normally open* relay is just the opposite. ("Normal" in this sense means no current in the coil.) The relay in the illustration (Fig. 8-8) can be used either as 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.

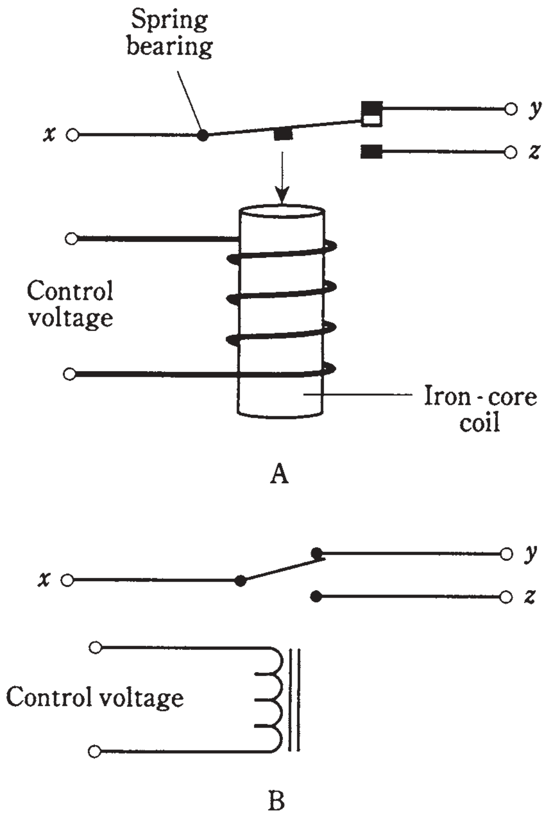
Some relays have several sets of contacts. Some relays are meant to remain in one state (either with current or without) for a long time, while others are meant to switch several times per second. The fastest relays work dozens of times per second. These are used for such purposes as keying radio transmitters in Morse code or radioteletype.

The dc motor

Magnetic fields can produce considerable mechanical forces. These forces can be harnessed to do work. The device that converts direct-current energy into rotating mechanical energy is a *dc motor.*

Motors can be microscopic in size, or as big as a house. Some tiny motors are being considered for use in medical devices that can actually circulate in the bloodstream or be installed in body organs. Others can pull a train at freeway speeds.

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. The greater the current that flows in the coils, the stronger the torque, and the more electrical energy is needed.



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

Figure 8-9 is a simplified, cutaway drawing of a dc motor. One set of coils, called the *armature coil,* goes around 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 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.

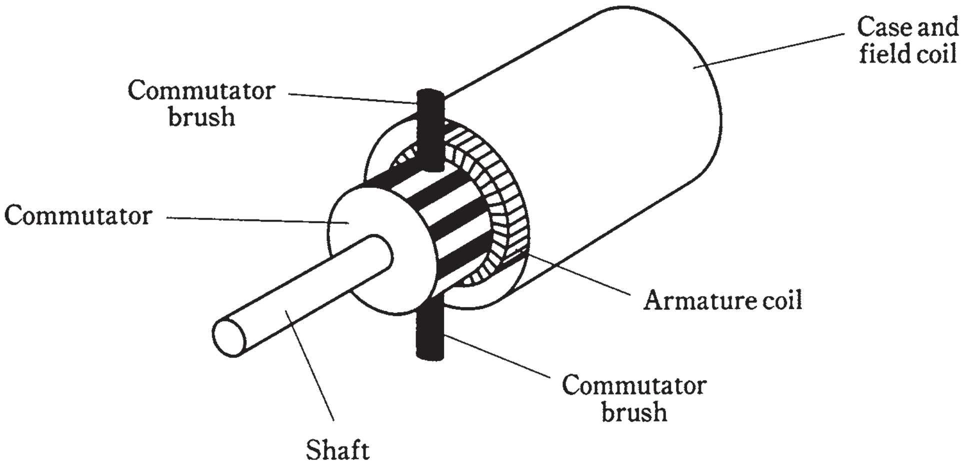
Some dc motors can also be used to generate direct current. These motors contain permanent magnets in place of one of the sets of coils. When the shaft is rotated, a pulsating direct current appears across the coil.

Magnetic data storage

Magnetic fields can be used to store data in different forms. Common media for data storage include the *magnetic tape,* the *magnetic disk,* and *magnetic bubble memory.*

Magnetic tape

Recording tape is the stuff you find in cassette players. It is also sometimes seen on reel-to-reel devices. These days, magnetic tape is used for home entertainment, especially hi-fi music and home video.



8-9 Cutaway view of a dc motor.

The tape itself consists of millions of particles of iron oxide, attached to a plastic or mylar strip. 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, controlled speed. This produces regions in which the iron-oxide 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, for different applications. The 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.

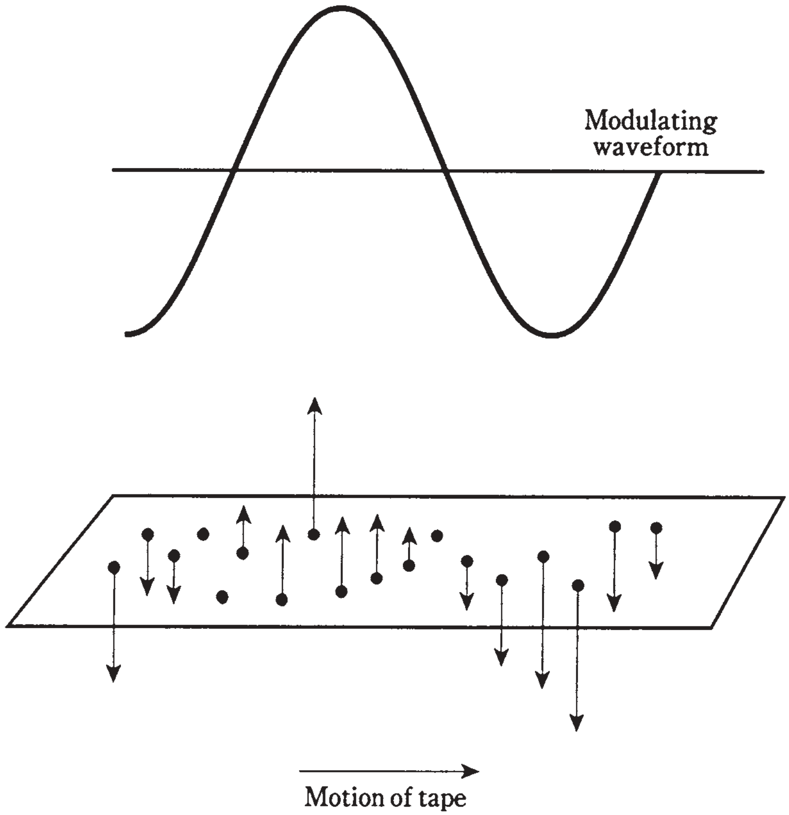
The data on a magnetic tape can be distorted or erased by external magnetic fields. Therefore, tapes should be protected from such fields. Keep magnetic tape away from magnets. Extreme heat can also result in loss of data, and possibly even physical damage to the tape.

Magnetic disk

The age of the personal computer has seen the development of ever-more-compact data-storage systems. One of the most versatile is the *magnetic disk.*

A magnetic disk can be either rigid or flexible. Disks are available in various sizes. *Hard disks* store the most data, and are generally found inside of computer units. *Floppy disks* or *diskettes* are usually either $3^1/_2$ or $5^1/_4$ inch in diameter, and can be inserted and removed from recording/playback machines called *disk drives.*

The principle of the magnetic disk, on the micro scale, is the same as that of the magnetic tape. The information is stored in *digital* form; that is, there are only two different ways that the particles are magnetized. This results in almost perfect, error-free storage.



8-10 On recording tape, particles are magnetized in a pattern that follows the modulating waveform.

On a larger scale, the disk works differently than the tape, simply 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. But on a disk, no two bits are ever farther apart than the diameter of the disk. This means that data can be stored and retrieved much more quickly onto, or from, a disk than is possible with a tape.

A typical diskette can store an amount of digital information equivalent to a short novel.

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

Magnetic 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. This type of memory is used in large computer systems, because it allows the storage, retrieval, and transfer of great quantities of data. The bits of data are stored as tiny magnetic fields, in a medium that is made from magnetic film and semiconductor materials.

A full description of the way bubble memory systems are made, and the way they work, is too advanced for this book. 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.

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. Is what makes a compass work.
   D. Is what 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 always slanted, no matter where you go.
   D. Are exactly symmetrical around the earth, even far out into space.

3. A material that can be permanently magnetized is generally said to be:
   A. Magnetic.
   B. Electromagnetic.
   C. Permanently magnetic.
   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. Magnetic flux can always be attributed to:
   A. Ferromagnetic materials.
   B. Aligned atoms.
   C. Motion of charged particles.
   D. The geomagnetic field.

6. Lines of magnetic flux are said to originate:
   A. In atoms of ferromagnetic materials.
   B. At a north magnetic pole.
   C. Where the lines converge to a point.
   D. In 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 in ampere-turns?
   A. 5000.
   B. 50.
   C. 5.0.
   D. 0.02.

11. If a wire coil has 100 turns and carries 1.30 A of current, what is the magnetomotive force in gilberts?
   A. 130.
   B. 76.9.
   C. 164.
   D. 61.0.

12. Which of the following is *not* generally possible in a geomagnetic storm?
   A. Charged particles streaming out from the sun.
   B. Fluctuations in the earth's magnetic field.
   C. Disruption of electrical power transmission.

D. Disruption of microwave radio links.

13. An ac electromagnet:

A. Will attract only other magnetized objects.

B. Will attract pure, unmagnetized iron.

C. Will repel other magnetized objects.

D. Will either attract or repel permanent magnets, depending on the polarity.

14. An advantage of an electromagnet over a permanent magnet is 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.

15. A substance with high retentivity is best suited for making:

A. An ac electromagnet.

B. A dc electromagnet.

C. An electrostatic shield.

D. A permanent magnet.

16. A relay is connected into a circuit so that a device gets a signal only when the relay coil carries current. The relay is probably:

A. An ac relay.

B. A dc relay.

C. Normally closed.

D. Normally open.

17. A device that reverses magnetic field polarity to keep a dc motor rotating is:

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. Voices.

B. Video.

C. Digital data.

D. All of the above.

19. An advantage of a magnetic disk, as 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 bubble memory is best suited for:

A. A large computer.

B. A home video entertainment system.

C. A portable cassette player.

D. A magnetic disk.

Test: Part one

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. A 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, is zero.
   - D. The net resistance of a parallel set of resistors is less than the value of the smallest resistor.
   - E. The total power 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.
   B. A wiring diagram shows component values.
   C. A schematic does not show all the interconnections between the components.
   D. A schematic shows pictures of components, while a wiring diagram shows the electronic symbols.
   E. A schematic shows the electronic symbols, while a wiring diagram shows pictures of the components.
5. Which of the following is a good use, or place, for a wirewound resistor?
   A. To dissipate a large amount of dc power.
   B. In the input of a radio-frequency amplifier.
   C. In the output of a radio-frequency amplifier.
   D. In an antenna, 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 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 going 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. A charged atom is known as:
   A. A molecule.
   B. An isotope.
   C. An ion.
   D. An electron.
   E. A fundamental particle.

13. A battery delivers 12 V to a bulb. The current in the bulb is 3 A. What is the resistance of the bulb?
   A. 36 Ω.
   B. 4 Ω.
   C. 0.25 Ω.
   D. 108 Ω.
   E. 0.75 Ω.

14. Peak values are always:
   A. Greater than average values.
   B. Less than average values.
   C. Greater than or equal to average values.
   D. Less than or equal to average values.

E. Fluctuating.

15. A resistor has a value of 680 ohms, and a tolerance of plus or minus 5 percent. Which of the following values 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 that they:

A. Pollute the environment when discarded.

B. Supply less voltage than other cells.

C. Can reverse polarity unexpectedly.

D. Must be physically large.

E. Must be kept right-side-up.

18. A battery supplies 6.0 V to a bulb rated at 12 W. How much current does the bulb draw?

A. 2.0 A.

B. 0.5 A.

C. 72 A.

D. 40 mA.

E. 72 mA.

19. Of the following, which is not a common use of a resistor?

A. Biasing for a transistor.

B. Voltage division.

C. Current limiting.

D. Use as a “dummy” antenna.

E. Increasing the charge in a capacitor.

20. When a charge builds up without a 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. Equal value, and the same polarity, as the supply.
    B. A value that depends on the ratio of the resistances.
    C. Different value from, but the same polarity as, the supply.
    D. Equal value as, but opposite polarity from, the supply.
    E. Different value, and 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, called its:
    A. Molecule.
    B. Electron.
    C. Proton.
    D. Atom.
    E. Isotope.

24. An advantage of a magnetic disk over magnetic tape for data storage is 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. A 6-V battery is connected across a series combination of resistors. The resistance values are 1, 2, and 3 Ω. What is the current through the 2-Ω resistor?
    A. 1 A.
    B. 3 A.
    C. 12 A.
    D. 24 A.
    E. 72 A.

26. A material that has extremely high electrical resistance 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 have exactly 1.500 V.
D. Cannot be recharged.
E. Are made of zinc and carbon.

28. A rheostat:
A. Is 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. A voltage typical of a dry cell is:
A. 12 V.
B. 6 V.
C. 1.5 V.
D. 117 V.
E. 0.15 V.

30. A geomagnetic storm:
A. Causes solar wind.
B. Causes charged particles to bombard the earth.
C. Can disrupt the earth's magnetic field.
D. Ruins microwave communications.
E. Has no effect near the earth's poles.

31. An advantage of an alkaline cell over a zinc-carbon cell is that:
A. The alkaline cell provides more voltage.
B. The alkaline cell can be recharged.
C. An alkaline cell works at lower temperatures.
D. The alkaline cell is far less bulky for the same amount of energy capacity.
E. There is no advantage of alkaline over zinc-carbon cells.

32. A battery delivers 12 V across a set of six 4-Ω resistors in a series voltage dividing combination. This provides six different voltages, differing by an increment of:
A. $\frac{1}{4}$ V.
B. $\frac{1}{3}$ V.

C. 1 V.

D. 2 V.

E. 3 V.

33. A unit of electrical charge quantity is the:

A. Volt.

B. Ampere.

C. Watt.

D. Tesla.

E. Coulomb.

34. A unit of sound volume is:

A. The volt per square meter.

B. The volt.

C. The watt hour.

D. The decibel.

E. The ampere per square meter.

35. A 24-V battery is connected across a set of four resistors in parallel. Each resistor has a value of 32 ohms. What is the total power dissipated by the resistors?

A. 0.19 W.

B. 3 W.

C. 192 W.

D. 0.33 W.

E. 72 W.

36. The main difference between a "lantern" battery and a "transistor" battery is:

A. The lantern battery has higher voltage.

B. The lantern battery has more energy capacity.

C. Lantern batteries cannot be used with electronic devices such as transistor radios.

D. Lantern batteries can be recharged, but transistor batteries cannot.

E. The lantern battery is more compact.

37. NICAD batteries are most extensively used:

A. In disposable flashlights.

B. In large lanterns.

C. As car batteries.

D. In handheld radio transceivers.

E. In remote garage-door-opener control boxes.

38. A voltmeter should have:

A. Very 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 a core with high retentivity.
C. Will 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. A 12-V battery is connected to a set of three resistors in series. The resistance values are 1,2, and 3 ohms. What is the voltage across the 3-Ω resistor?
A. 1 V.
B. 2 V.
C. 4 V.
D. 6 V.
E. 12 V.

43. Nine 90-ohm resistors are connected in a 3×3 series-parallel network. The total resistance is:
A. 10 Ω.
B. 30 Ω.
C. 90 Ω.
D. 270 Ω.
E. 810 Ω.

44. A device commonly used for remote switching of wire communications signals is:
    - A. A solenoid.
    - B. An electromagnet.
    - C. A potentiometer.
    - D. A photovoltaic cell.
    - E. A relay.

45. NICAD memory:
    - A. Occurs often when NICADs are misused.
    - B. Indicates that the cell or battery is dead.
    - C. Does not occur very often.
    - D. Can cause a NICAD to explode.
    - E. Causes NICADs to reverse polarity.

46. A 100-W bulb burns for 100 hours. It has consumed:
    - 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. 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 a bar-graph meter is that:

A. Is isn't very sensitive.

B. It isn't stable.

C. It can't give a very precise reading.

D. You need special training to read it.

E. It shows only peak values.

2
PART
Alternating current

This page intentionally left blank

9
CHAPTER

Alternating current basics

DIRECT CURRENT (DC) IS SIMPLE. IT CAN BE EXPRESSED IN TERMS OF JUST TWO variables: polarity (or direction), and magnitude. Alternating current (ac) is somewhat more complicated, because there are three things that can vary. Because of the greater number of parameters, alternating-current circuits behave in more complex ways than direct-current circuits. This chapter will acquaint you with the most common forms of alternating current. A few of the less often-seen types are also mentioned.

Definition of alternating current

Recall that direct current has a polarity, or direction, that stays the same over a long period of time. Although the magnitude might vary—the number of amperes, volts, or watts can fluctuate—the charge carriers always flow in the same direction through the circuit.

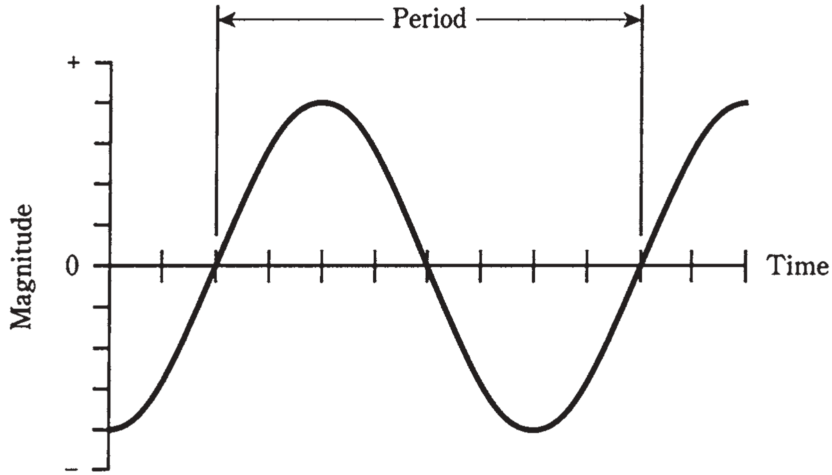
In alternating current, the polarity reverses again and again at regular intervals. The magnitude usually changes because of this constant reversal of polarity, although there are certain cases where the magnitude doesn't change even though the polarity does.

The rate of change of polarity is the third 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 *magnitude versus time* repeats itself over and over, so that the same pattern recurs countless times. 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. Radio-frequency currents reverse polarity millions or billions of times a second. The charged particles held captive by the magnetic field of the sun, and perhaps also by the much larger magnetic fields around galaxies, might reverse their direction over periods measured in thousands or millions of years. Period, when measured in seconds, is denoted by *T*.



9-1 A sine wave. The period is the length of time for one complete cycle.

The *frequency,* denoted *f,* of a wave is the reciprocal of the period. That is, $f = 1/T$ and $T = 1/f$. Originally, frequency was specified in *cycles per second,* abbreviated *cps*. High frequencies were sometimes given in *kilocycles, megacycles,* or *gigacycles,* representing thousands, millions, or billions of cycles per second. But nowadays, the unit is known as the *hertz,* abbreviated 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:

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

Sometimes an even bigger unit, the *terahertz (THz),* is needed. This is a trillion (1,000,000,000,000) hertz. Electrical currents generally do not attain such frequencies, although *electromagnetic radiation* can.

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 might also be components at "odd" frequencies. Some ac waves are extremely complex, consisting of hundreds, thousands, or even infinitely many different component, frequencies. In this book, most of the attention will be given to ac waves that have just one frequency.

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 will have a perfect sine waveshape. And any perfect sine-wave current contains 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 has an almost perfect sine waveshape, with a frequency of 60 Hz.

The square wave

Earlier in this chapter, it was said that there can be an alternating current whose magnitude never changes. You might at first think this is impossible. How can polarity reverse without some change in the level? The *square wave* is an example of this.

On an oscilloscope, a perfect square wave looks like a pair of parallel, dotted 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, versus time on the horizontal scale. The transitions between negative and positive for a true square wave don't show up on the oscilloscope, because they're instantaneous. But perfection is rare. Usually, the transitions can be seen as vertical lines (Fig. 9-2B).

A square wave might have equal negative and positive peaks. Then the absolute magnitude of the wave is constant, at a certain voltage, current, or power level. Half of the time it's +x, and the other half it's -x volts, amperes, or watts. Some square waves are lopsided, with the positive and negative magnitudes differing.

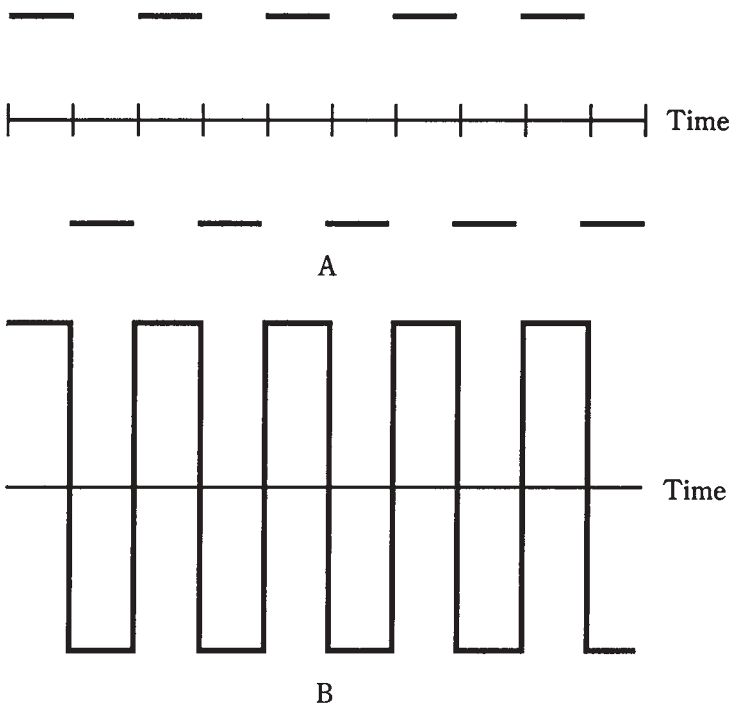
Sawtooth waves

Some ac waves rise and fall in straight 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.

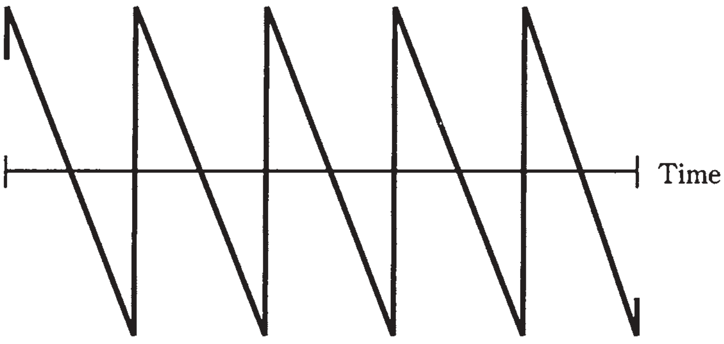
Sawtooth waves are generated by certain electronic test devices. These waves provide ideal signals for control purposes. Integrated circuits can be wired so that they produce sawtooth waves having an exact desired shape.

Fast-rise, slow-decay

In Fig. 9-3, one form of sawtooth wave is shown. The positive-going slope (rise) is extremely steep, as with a square wave, but the negative-going slope (fall or decay) is gradual. The period of the wave is the time between points at identical positions on two successive pulses.



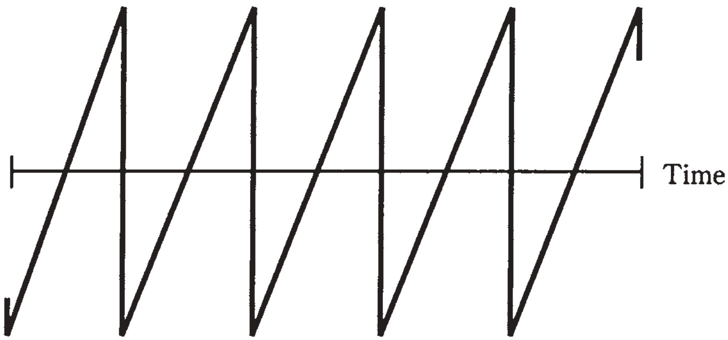
9-2 At A, a perfect square wave. At B, the more common rendition of a square wave.



9-3 Fast-rise, slow-decay sawtooth.

Slow-rise, fast-decay

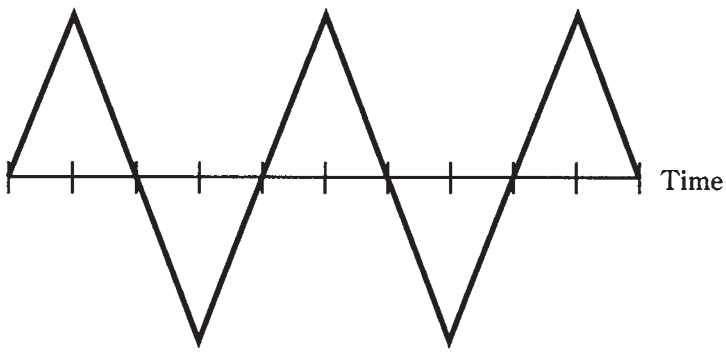
Another form of sawtooth wave is just the opposite, with a gradual positive-going slope and a vertical negative-going transition. This type of wave is sometimes called a *ramp*, because it looks like an incline going upwards (Fig. 9-4). This waveshape is useful for scanning in television sets and oscilloscopes. It tells the electron beam to move, or trace, at a constant speed from left to right during the upwards sloping part of the wave. Then it retraces, or brings the electron beam back, at a high speed for the next trace.



9-4 Slow-rise, fast-decay sawtooth, also called a ramp.

Variable rise and decay

You can probably guess that sawtooth waves can have rise and decay slopes in an infinite number of different combinations. One example is shown in Fig. 9-5. In this case, the positive-going slope is the same as the negative-going slope. This is a *triangular wave.*



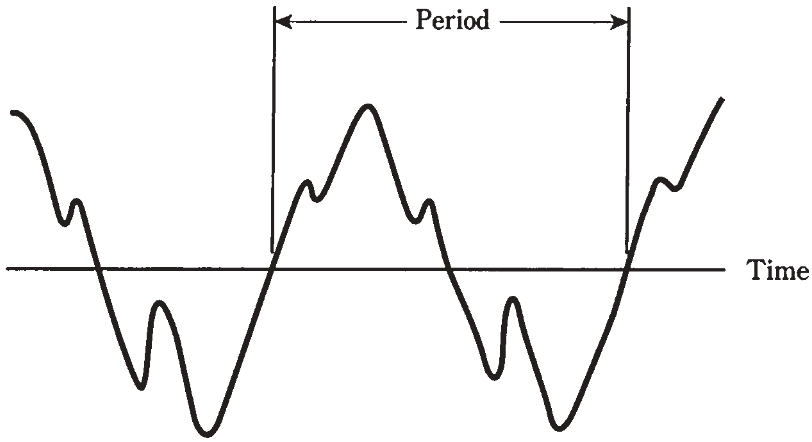
9-5 Triangular wave.

Complex and irregular waveforms

The shape of an ac wave can get exceedingly complicated, but as long as it has a definite period, and as long as the polarity keeps switching back and forth between positive and negative, it is ac.

Fig. 9-6 shows an example of a complex ac wave. You can see that there is a 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 tell the period. This is because the wave has two or more components that are nearly the same magnitude. When this happens, the *frequency spectrum* of the wave will be multifaceted. The energy is split up among two or more frequencies.



9-6 An irregular waveform.

Frequency spectrum

An oscilloscope shows a graph of magnitude versus time. Because time is on the horizontal axis, the oscilloscope is said to be a *time-domain* instrument.

Sometimes you want to see magnitude 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 with a cathode-ray display similar to an oscilloscope. Its horizontal axis shows frequency, from some adjustable minimum (extreme left) to some adjustable maximum (extreme right).

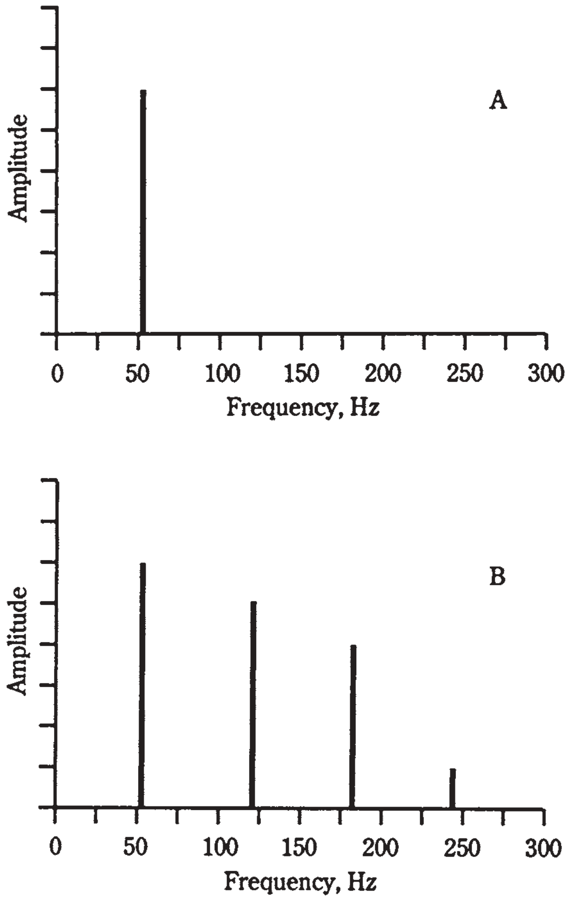
An ac sine wave, as displayed on a spectrum analyzer, appears as a single "pip, " or vertical line (Fig. 9-7A). This means that all of the energy in the wave is concentrated at one single frequency.

Many ac waves contain *harmonic* energy along with the fundamental, or main, frequency. A harmonic frequency is a whole-number multiple of the fundamental frequency. For example, if 60 Hz is the fundamental, 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.*

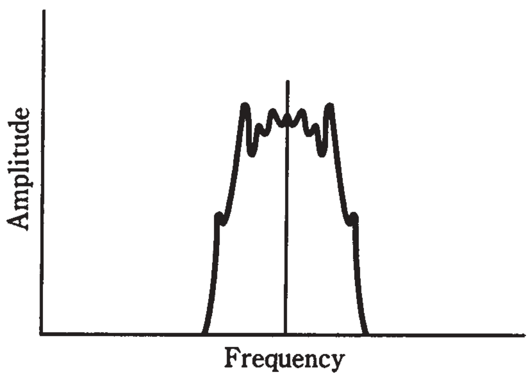
In general, if a wave has a frequency equal to n times the fundamental, then that wave is the *nth harmonic.* In the illustration of Fig. 9-7B, a wave is shown along with several harmonics, as it would look on the display screen of a spectrum analyzer.

The frequency spectra of square waves and sawtooth waves contain harmonic energy in addition to the fundamental. The wave shape depends on the amount of energy in the harmonics, and the way in which this energy is distributed among the harmonic frequencies. A detailed discussion of these relationships is far too sophisticated for this book.

Irregular waves can have practically any imaginable frequency distribution. An example is shown at Fig. 9-8. This is a display of a voice-modulated radio signal. Much of the energy is concentrated at the center of the pattern, at the frequency shown by the vertical line. But there is also plenty of energy "splattered around" this *carrier* frequency. On an oscilloscope, this signal would look like a fuzzy sine wave, indicating that it is ac, although it contains a potpourri of minor components.



9-7 At A, pure 60-Hz sine wave on spectrum analyzer. At B, 60-Hz wave containing harmonics.



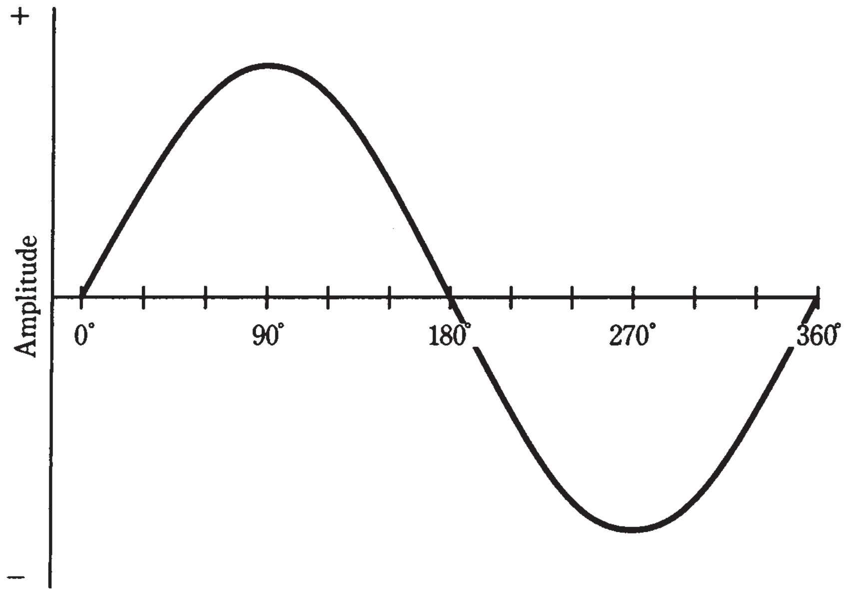
9-8 Modulated radio signal on spectrum analyzer

Little bits of a cycle

Engineers break the ac cycle down into small parts for analysis and reference. One complete cycle can be likened 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.* The value 0 degrees is assigned to the point in the cycle where the magnitude is 0 and positive-going. The same point on the next cycle is given the value 360 degrees. Then halfway through the cycle is 180 degrees; a quarter cycle is 90 degrees, and so on. This is illustrated in Fig. 9-9.



9-9 A cycle is divided into 360 degrees.

Radians

The other method of specifying phase is to divide the cycle into 6.28 equal parts. This is approximately the number of radii of a circle that can be laid end-to-end around the circumference. A *radian* of phase is equal to about 57.3 degrees. This unit of phase is something you won't often be needing to use, because it's more common among physicists than among engineers.

Sometimes, the frequency of an ac wave is measured in radians per second, rather than in hertz (cycles per second). Because there are about 6.28 radians in a complete cycle of 360 degrees, the *angular frequency* of a wave, in radians per second, is equal to about 6.28 times the frequency in hertz.

Phase difference

Two ac waves might have exactly the same frequency, but they can still 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,* signal.

If two ac waves have the same frequency and the same magnitude, but differ in phase by 180 degrees (a half cycle), they will cancel each other out, and the net signal will be zero. If the two waves are in phase, the resulting signal will have the same frequency, but twice the amplitude of either signal alone.

If two ac waves have the same frequency but different magnitudes, and differ in phase by 180 degrees, the resulting composite signal will have the same frequency as the originals, and a magnitude equal to the difference between the two. If two such waves are exactly in phase, the composite will have the same frequency as the originals, and a magnitude equal to the sum of the two.

If the waves have the same frequency but differ in phase by some odd amount such as 75 degrees or 310 degrees, the resulting signal will have the same frequency, but will not have the same waveshape as either of the original signals. The variety of such cases is infinite.

Household utility current, as you get it from wall outlets, consists of a 60-Hz sine wave with just one phase component. But the energy is transmitted over long distances in three phases, each differing by 120 degrees or 1/3 cycle. This is what is meant by *three-phase ac*. Each of the three ac waves carries 1/3 of the total power in a utility transmission line.

Amplitude of alternating current

Amplitude is sometimes called magnitude, level, or intensity. Depending on the quantity being measured, the magnitude of an ac wave might be given in amperes (for current), volts (for voltage), or watts (for power).

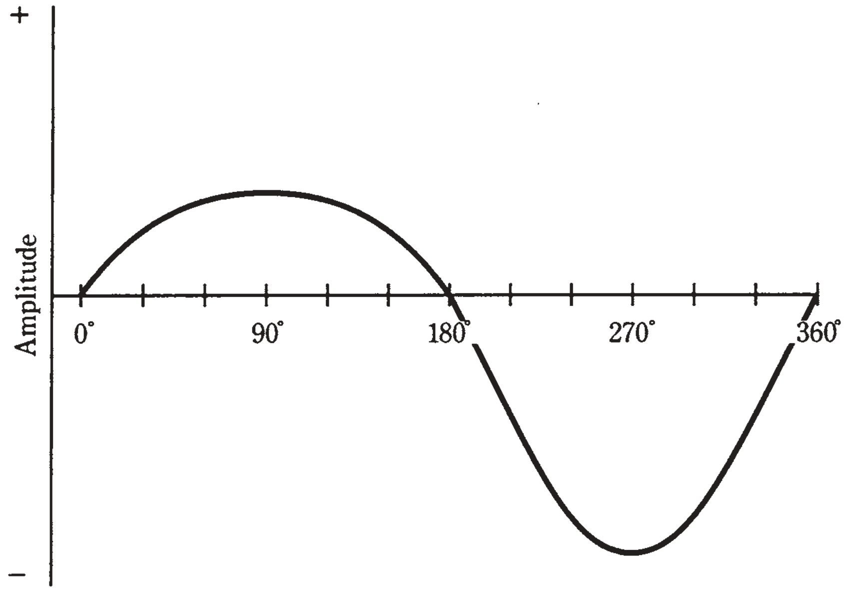
Instantaneous amplitude

The *instantaneous amplitude* of an ac wave is the amplitude at some precise moment in time. This constantly changes. The manner in which it varies depends on the waveform. You have already seen renditions of common ac waveforms in this chapter. Instantaneous amplitudes are represented by individual points on the wave curves.

Peak amplitude

The peak amplitude of an ac wave is the maximum extent, either positive or negative, that the instantaneous amplitude attains.

In many waves, the positive and negative peak amplitudes 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.



9-10 A wave with unequal positive and negative peaks.

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). Another way of saying this is that the pk-pk amplitude is equal to the positive peak amplitude plus the negative peak amplitude. Peak-to-peak is a way of expressing how much the wave level "swings" during the cycle.

In many waves, the pk-pk amplitude is just twice the peak amplitude. This is the case when the positive and negative peak amplitudes are the same.

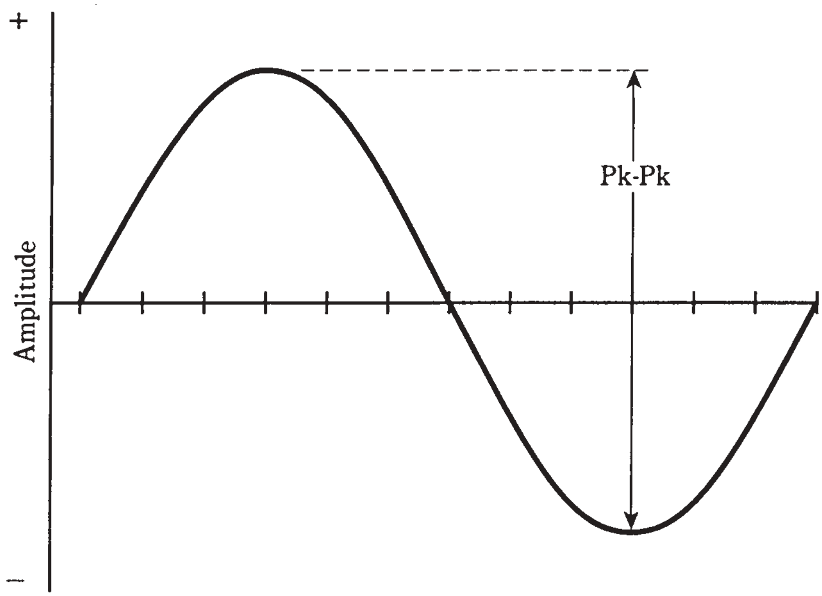
Root-mean-square amplitude

Often, it is necessary to express the effective level 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. When you say a wall outlet has 117 V, you mean 117 effective volts. The most common figure for effective ac levels is called the *root-mean-square,* or rms, value.

For a perfect sine wave, the rms value is equal to 0.707 times the peak value, or 0.354 times the pk-pk value. Conversely, the peak value is 1.414 times the rms value, and the pk-pk value is 2.828 times the rms value. The rms figures are most often used with perfect sine waves, such as the utility voltage, or the effective voltage of a radio signal.

For a perfect square wave, the rms value is the same as the peak value. The pk-pk value is twice the rms value or the peak value.

For sawtooth and irregular waves, the relationship between the rms value and the peak value depends on the shape of the wave. But the rms value is never more than the peak value for any waveshape.



9-11 Peak-to-peak amplitude.

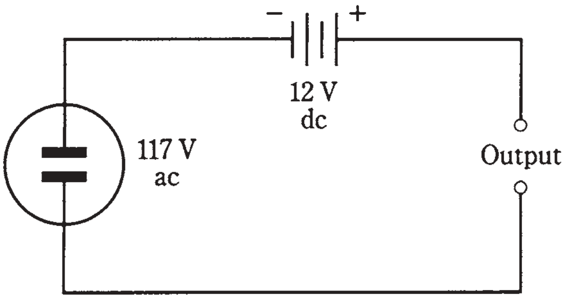
The name "root mean square" was not chosen just because it sounds interesting. It literally means that the value of a wave is mathematically operated on, by taking the square root of the mean of the square of all its values. You don't really have to be concerned with this process, but it's a good idea to remember the above numbers for the relationships between peak, pk-pk, and rms values for sine waves and square waves.

For 117 V rms at a utility outlet, the peak voltage is considerably greater. The pk-pk voltage is far greater.

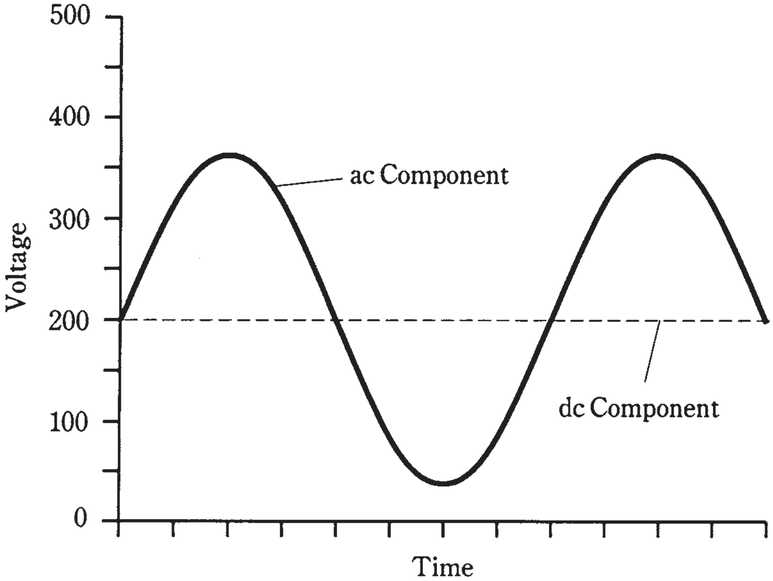
Superimposed direct current

Sometimes a wave can have components of both ac and dc. The simplest example of an ac/dc combination is illustrated by the connection of a dc source, such as a battery, in series with an ac 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!) Then the ac wave will be displaced either positively or negatively by 12 V, depending on the polarity of the battery. This will result in a sine wave at the output, but one peak will be 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-Vdc source were connected in series with the utility output. 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 illustrated by Fig. 9-13.



9-12 Connection of a dc source in series with an ac source.



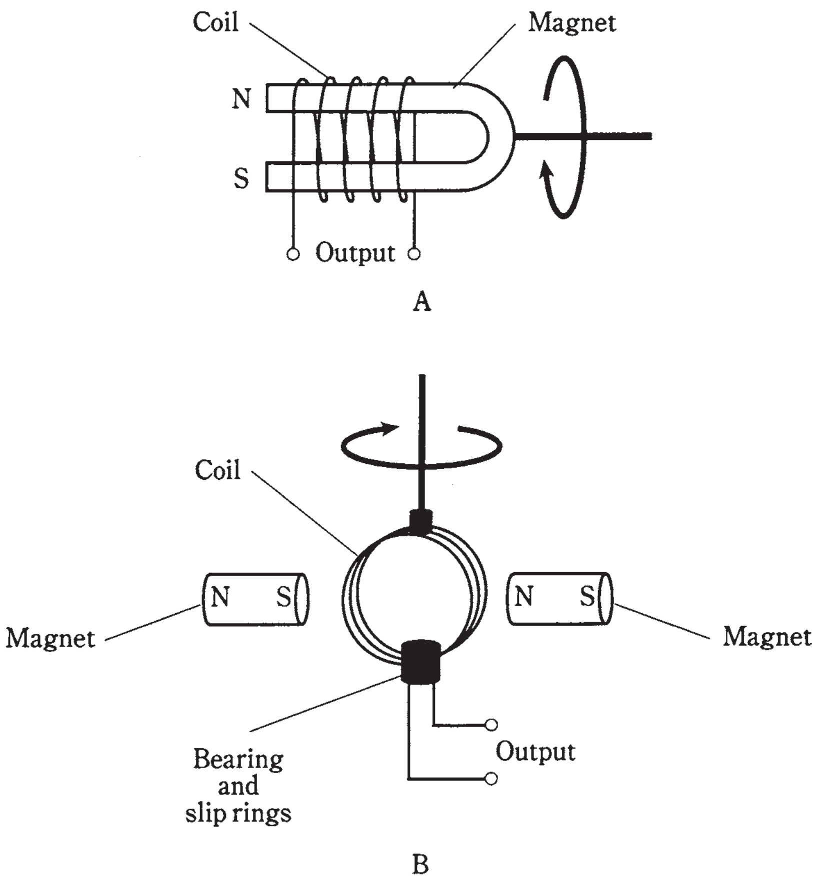
9-13 Waveform resulting from 117 Vac in series with + 200 Vdc.

"Hybrid" ac/dc combinations are not often generated deliberately. But such waveforms are sometimes seen at certain points in electronic circuitry.

The ac generator

Alternating current is easily generated by means of a rotating magnet in a coil of wire (Fig. 9-14A), or by a rotating coil of wire inside a powerful magnet (Fig. 9-14B). In either case, the ac appears between the ends of the length of wire.

The ac voltage that a generator can develop 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 rotation. Normally, for utility ac, this speed is 3,600 revolutions per minute (rpm), or 60 complete revolutions per second (rps), so that the frequency is 60 Hz.



9-14 Two forms of ac generator. At A, the magnet rotates; at B, the coil rotates.

When a load, such as a light bulb or heater, is connected to an ac generator, it becomes more difficult to turn the generator. The more power needed from a generator, the greater the amount of power required to drive it. This is why it is not possible to connect a generator to, for instance, your 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 50 W of power to run a small radio, but nowhere near enough to provide electricity for a household.

The efficiency of a generator is the ratio of the power output to the 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 to this ideal.

At power plants, the generators are huge. Each one is as big as a house. The generators are driven by massive turbines. The turbines are turned by various natural sources of energy. Often, steam drives the turbines, and the steam is obtained via heat derived from the natural energy source.

Why ac?

You might wonder why ac is even used. Isn't it a lot more complicated than dc?

Well, ac is easy to generate from turbines, as you've just seen. Rotating coil-and magnet devices always produce ac, and in order to get dc from this, *rectification* and *filtering* are necessary. These processes can be difficult to achieve with high voltages.

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. To serve millions of consumers, the immense power of falling or flowing water, the ocean tides, wind, burning fossil fuels, safe nuclear fusion, or of geothermal heat are needed. (Nuclear fission will work, but it is under scrutiny nowadays because it produces dangerous radioactive by-products.) 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.

Thomas Edison is said to have favored dc over ac for electrical power transmission in the early days, as utilities were first being planned. His colleagues argued that ac would work better. It took awhile to convince Mr. Edison to change his mind. He eventually did. But perhaps he knew something that his contemporaries did not foresee.

There is one advantage to direct current in utility applications. This is for 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's original sentiments will be at least partly vindicated. His was a long view.

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 not with dc?

   A. Power.

   B. Voltage.

   C. Frequency.

   D. Magnitude.

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, a pure ac signal, having just one frequency component,would look like:

   A. A single pip.
   B. A perfect sine wave.
   C. A square wave.
   D. A sawtooth wave.

4. The period of an ac wave is:

   A. The same as the frequency.
   B. Not related to the frequency.
   C. Equal to 1 divided by the frequency.
   D. Equal to the amplitude divided by the frequency.

5. The sixth harmonic of an ac wave whose period is 0.001 second has a frequency of

   A. 0.006 Hz.
   B. 167 Hz.
   C. 7 kHz.
   D. 6 kHz.

6. A degree of phase represents:

   A. 6.28 cycles.
   B. 57.3 cycles.
   C. 1/6.28 cycle.
   D. 1/360 cycle.

7. Two waves have the same frequency but differ in phase by 1/20 cycle. The phase difference in degrees is:

   A. 18.
   B. 20.
   C. 36.
   D. 5.73.

8. A signal has a frequency of 1770 Hz. The angular frequency is:

   A. 1770 radians per second.

B. 11,120 radians per second.

C. 282 radians per second.

D. Impossible to determine from the data given.

9. A triangular wave:

A. Has a fast rise time and a slow decay time.

B. Has a slow rise time and a fast decay time.

C. Has equal rise and decay rates.

D. Rises and falls abruptly.

10. Three-phase ac:

A. Has waves that add up to three times the originals.

B. Has three waves, all of the same magnitude.

C. Is what you get at a common wall outlet.

D. Is of interest only to physicists.

11. If two waves have the same frequency and the same amplitude, but opposite phase, the composite wave is:

A. Twice the amplitude of either wave alone.

B. Half the amplitude of either wave alone.

C. A complex waveform, but with the same frequency as the originals.

D. Zero.

12. If two waves have the same frequency and the same phase, the composite wave:

A. Has a magnitude equal to the difference between the two originals.

B. Has a magnitude equal to the sum of the two originals.

C. Is complex, with the same frequency as the originals.

D. Is zero.

13. In a 117-V utility circuit, the peak voltage is:

A. 82.7 V.

B. 165 V.

C. 234 V.

D. 331 V.

14. In a 117-V utility circuit, the pk-pk voltage is:

A. 82.7 V.

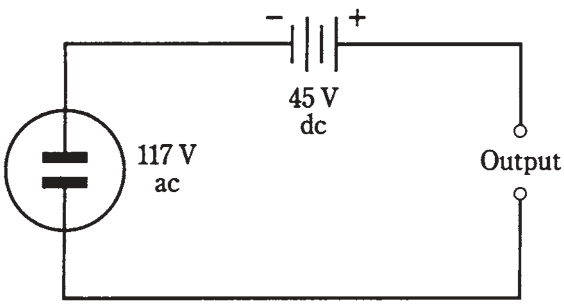
B. 165 V.

C. 234 V.

D. 331 V.

15. In a perfect sine wave, the pk-pk value is:
    A. Half the peak value.
    B. The same as the peak value.
    C. 1.414 times the peak value.
    D. Twice the peak value.

16. If a 45-Vdc battery is connected in series with the 117-V utility mains as shown in Fig. 9-15, the peak voltages will be:
    A. + 210 V and − 120 V.
    B. + 162 V and − 72 V.
    C. + 396 V and − 286 V.
    D. Both equal to 117 V.



9-15 Illustration for quiz question 16.

17. In the situation of question 16, the pk-pk voltage will be:
    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.
    B. Smooth ac.

C. Ac with one peak greater than the other.

D Pulsating dc.

20. An advantage of ac over dc in utility applications is:

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 gotten from dc generators.

D. Ac can be generated with less dangerous by-products.

10
CHAPTER
Inductance

THIS CHAPTER DELVES INTO DEVICES THAT OPPOSE THE FLOW OF AC BY temporarily storing some of the electrical energy as a magnetic field. Such devices are called inductors. The action of these components 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. Some active electronic devices display inductance, even when you don't think of the circuit in those terms.

Inductance can appear where it isn't wanted. Noncoil inductance becomes increasingly common as the frequency of an altemating current increases. At very-high, ultra-high, and microwave radio frequencies, this phenomenon becomes a major consideration in the design of communications equipment.

The property of inductance

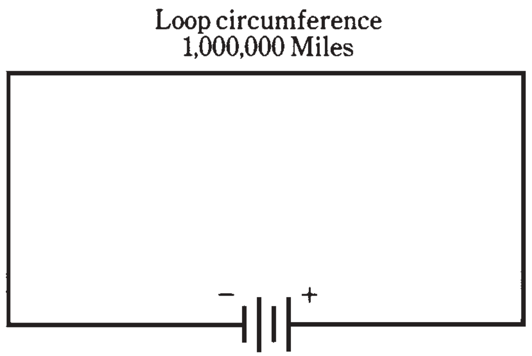
Suppose you have a wire a million miles long. What will happen if you make this wire into a huge loop, and connect its ends to the terminals of a battery (Fig. 10-1)?

You can surmise that a 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 will take a while for the electrons from the negative terminal to work their way around the loop to the positive terminal.

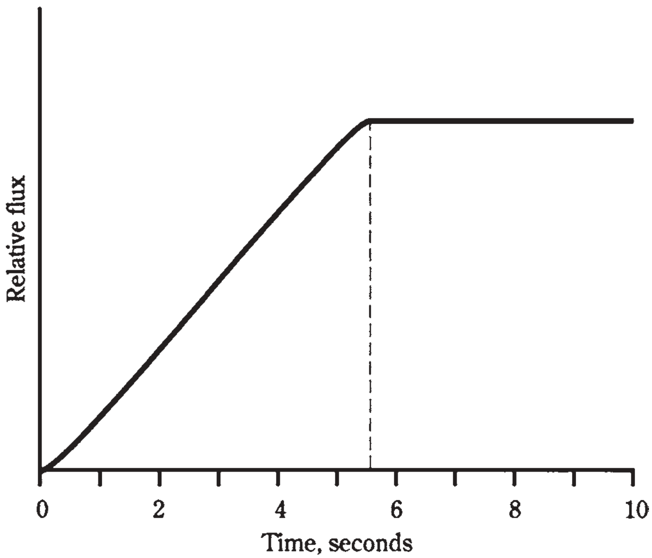
The effect of the current moves along the wire at a little less than the speed of light. In this case, it's about 180,000 miles per second, perhaps 97 percent of the speed of light in free space. It will take a little time for the current to build up to its maximum level. The first electrons won't start to enter the positive terminal until more than five seconds have passed.

The magnetic field produced by the loop will be small at first, because current is



10-1 A huge loop of wire illustrates the principle of inductance. See text.

flowing in only part of the loop. The flux will increase over a period of a few seconds, as the electrons get around the loop. Figure 10-2 is an approximate graph of the overall magnetic field versus time. After about 5.5 seconds, current is flowing around the whole loop, and the magnetic field has reached its maximum.



10-2 Relative magnetic flux in the huge wire loop, as a function of time in seconds.

A certain amount of energy is stored in this magnetic field. The ability of the loop to store energy in this way is the property of inductance. It is abbreviated by the letter *L*.

Practical inductors

Of course, it's not easy to make wire loops even approaching a million miles in circumference. But lengths of wire can be coiled up. When this is done, the magnetic flux is increased many times for a given length of wire compared with the flux produced by a single-turn loop. This is how inductors are made in practical electrical and electronic devices.

For any coil, the magnetic flux density is multiplied when a ferromagnetic core is placed within the coil of wire. Remember this from the study of magnetism. The increase in flux density has the effect of multiplying the inductance of a coil, so that it is many times greater with a ferromagnetic core than with an air core.

The current that an inductor can handle depends on the size of the wire. The inductance does not; it is a function of the number of turns in the coil, the diameter of the coil, and the overall shape of the coil.

In general, inductance of a coil is directly proportional to the number of turns of wire. Inductance is also 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 also: the longer the coil, the less the inductance.

The unit of inductance

When a battery is connected across a wire-coil inductor (or any kind of inductor), it takes a while for the current flow to establish itself throughout the inductor. The current changes at a rate that depends on the inductance: the greater the inductance, the slower the rate of change of current for a given battery voltage.

The unit of inductance is an expression of the ratio between the rate of current change and the voltage across an inductor. An inductance of one *henry,* abbreviated *H*, represents a potential difference of one volt across an inductor within which the current is increasing or decreasing at one ampere per second.

The henry is an extremely large unit of inductance. Rarely will you see an inductor anywhere near this large, although some power-supply filter chokes have inductances up to several henrys. Usually, inductances are expressed in *millihenrys (mH), microhenrys (μH),* or even in *nanohenrys (nH).* You should know your prefix multipliers fairly well by now, but in case you've forgotten, 1 mH = 0.001 H = 10^{-3} H, 1 μH = 0.001 mH = 0.000001 H = 10^{-6} H, and 1 nH = 0.001 μH = 10^{-9} H.

Very small coils, with few turns of wire, produce small inductances, in which the current changes quickly and the voltages are small. Huge coils with ferromagnetic cores, and having many turns of wire, have large inductances, in which the current changes slowly and the voltages are large.

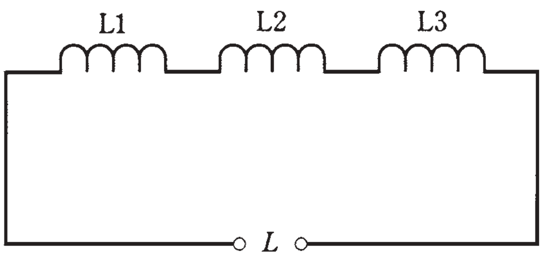
Inductors in series

As long as 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.

Problem 10-1

Three 40-μH inductors are connected in series, and there is no interaction, or *mutual inductances,* among them (Fig. 10-3). What is the total inductance?

You can just add up the values. Call the inductances of the individual components L_1, L_2, and L_3, and the total inductance L. Then $L = L_1 + L_2 + L_3 = 40 + 40 + 40 = 120\ \mu\text{H}$.



10-3 Inductors in series.

Problem 10-2

Suppose there are three inductors, with no mutual inductance, and their values are 20.0 mH, 55.0 μH, and 400 nH. What is the total inductance of these components if they are connected in series as shown in Fig. 10-3?

First, convert the inductances to the same units. You might use microhenrys, because that's the "middle-sized" unit here. Call $L_1 = 20.0$ mH $= 20{,}000\ \mu$H; $L_2 = 55.0\ \mu$H; $L_3 = 400$ nH $= 0.400$ uH. Then the total inductance is $L = 20{,}000 + 55.0 + 0.400$ uH $= 20{,}055.4\ \mu$H. The values of the original separate components were each given to three significant figures, so you should round the final figure off to 20,100 μH.

Note that subscripts are now used in designators. An example is L_2 (rather than $L2$). Some engineers like subscripts, while others don't want to bother with them. You should get used to seeing them both ways. They're both alright. If there are several inductors in series, and one of them has a value much larger than the values of the others, then the total inductance will be only a little bit more than the value of the largest inductor.

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, ..., L_n all connected in parallel. Then you can find the reciprocal of the total inductance, $1/L$, using the following formula:

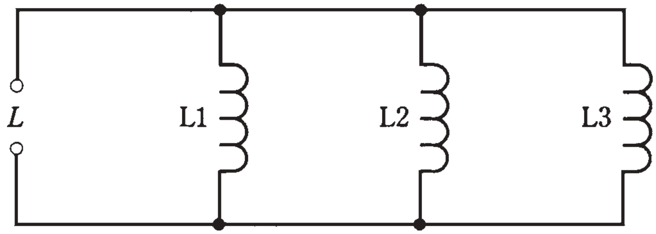
$$1/L = 1/L_1 + 1/L_2 + 1/L_3 + \ldots + 1/L_n$$

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. Don't mix microhenrys with millihenrys, or henrys with nanohenrys. The units you use for the individual component values will be the units you get for the final answer.

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 set?



10-4 Inductors in parallel.

Call the inductances $L_1 = 40$ μH, $L_2 = 40$ μH, and $L_3 = 40$ μH. Use the formula above 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 that there are four inductors in parallel, with no mutual inductance. Their values are $L_1 = 75$ mH, $L_2 = 40$ mH, $L_3 = 333$ μH, and $L_4 = 7.0$ H. What is the total net inductance?

You can use henrys, millihenrys, or microhenrys as the standard units in this problem. Suppose you decide to use henrys. Then $L_1 = 0.075$ H, $L_2 = 0.040$ H, $L_3 = 0.000333$ H, and $L_4 = 7.0$ H. Use the above formula to obtain $1/L = 13.33 + 25 + 3003 + 0.143 = 3041.473$. The reciprocal of this is the inductance $L = 0.00032879$ H $= 328.79$ μH. This can be rounded off to 330 μH because of significant-digits considerations.

This is just about the same as the 333-μH inductor alone. In real life, you could have only the single 333-μH inductor in this circuit, and the inductance would be essentially the same as with all four inductors.

If there are several inductors in parallel, and one of them has a value that is far smaller than the values of all the others, then the total inductance is just a little smaller than the value of the smallest inductor.

Interaction among inductors

In practical electrical circuits, there is almost always some mutual inductance between or among coils when they are wound in a cylindrical shape. The magnetic fields extend significantly outside solenoidal coils, and mutual effects are almost inevitable. The same is true between and among lengths of wire, especially at very-high, ultra-high, and microwave radio frequencies. Sometimes, mutual inductance is all right, and doesn't have a detrimental effect on the behavior of a circuit. But it can be a bad thing.

Mutual inductance can be minimized by using *shielded* wires and *toroidal* inductors. The most common shielded wire is *coaxial cable.* Toroidal inductors are discussed a little later in this chapter.

Coefficient of coupling

The *coefficient of coupling,* specified by the letter k, is a number ranging from 0 (no coupling) to 1 (maximum possible coupling). Two coils that are separated by a sheet of solid iron would have essentially $k = 0$; two coils wound on the same form, one right over the other, would have practically $k = 1$.

Mutual inductance

The mutual inductance is specified by the letter 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.

For two inductors, having values of L_1 and L_2 (both expressed in the same size units), and with a coefficient of coupling 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,

$$M = k\ (L_1 L_2)^{1/2}$$

Effects of mutual inductance

Mutual inductance can operate either to increase the inductance of a pair of series connected inductors, or to decrease it. This is because the magnetic fields might reinforce each other, or they might act against each other.

When two inductors are connected in series, and there is *reinforcing* mutual inductance between them, the total inductance L is given in the formula:

$$L = L_1 + L_2 + 2M$$

where L_1 and L_2 are the values of the individual inductors, and M is the mutual inductance. All inductances must be expressed in the same size units.

Problem 10-5

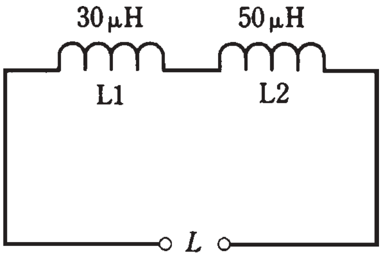
Suppose two coils, having values of 30 μH and 50 μH, are connected in series so that their fields reinforce, as shown in Fig. 10-5, and that the coefficient of coupling is 0.5. What is the total inductance of the combination?

First, calculate M from k. According to the formula for this, given above, $M = .5(50 \times 30)^{1/2} = 19.4\ \mu$H. Then the total inductance is equal to $L = L_1 + L_2 + 2M = 30 + 50 + 38.8 = 118.8\ \mu$H, rounded to 120 μH because only two significant digits are justified.

When two inductors are connected in series and the mutual inductance is in *opposition*, the total inductance L is given by the formula

$$L = L_1 + L_2 - 2M$$

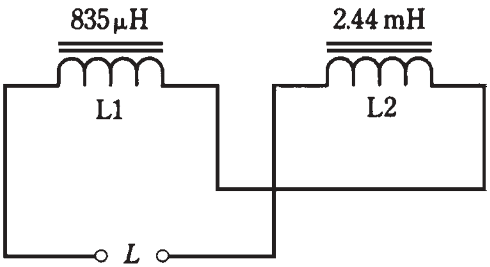
where, again, L_1 and L_2 are the values of the individual inductors.



10-5 Illustration for Problem 10-5.

Problem 10-6

There are two coils with values $L_1 = 835\ \mu$H and $L_2 = 2.44$ mH. 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. Notice that the coil inductances are specified in different units. Convert them both to microhenrys, so that L_2 becomes 2440 μH. Then $M = 0.922(835 \times 2440)^{1/2} = 1316\ \mu$H. The total inductance is therefore $L = L_1 + L_2 - 2M = 835 + 2440 - 2632 = 643\ \mu$H.

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 cancel each other. If two equal-valued inductors are connected in series so that their fluxes oppose, the result will be theoretically zero inductance.

Air-core coils

The simplest inductors (besides plain, straight lengths of wire) are coils. A coil can be wound on a plastic, wooden or other nonferromagnetic material, and it will work very well, although no *air-core* inductor can have very much inductance. In practice, the maximum attainable inductance for such coils is about 1 mH.

Air-core coils are used mostly at radio frequencies, in transmitters, receivers, and antenna networks. In general, the higher the frequency of an alternating current, the less inductance is needed to produce significant effects. Air-core coils can be made to have almost unlimited current-carrying capacity, just 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 is almost lossless. For these reasons, air-core coils can be made highly efficient.

Powdered-iron and ferrite 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 hundreds, thousands, and even millions of times. A small coil can thus be made to have a large inductance.

Powdered-iron cores are common at radio frequencies. *Ferrite* has a higher permeability than powdered iron, causing a greater concentration of magnetic flux lines within the coil. Ferrite is used at lower radio frequencies and at audio frequencies, as well as at medium and high radio frequencies.

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. 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 waste considerable power as heat. If a core gets hot enough, it might fracture. This will permanently change the inductance of the coil, and will also reduce its current-handling ability.

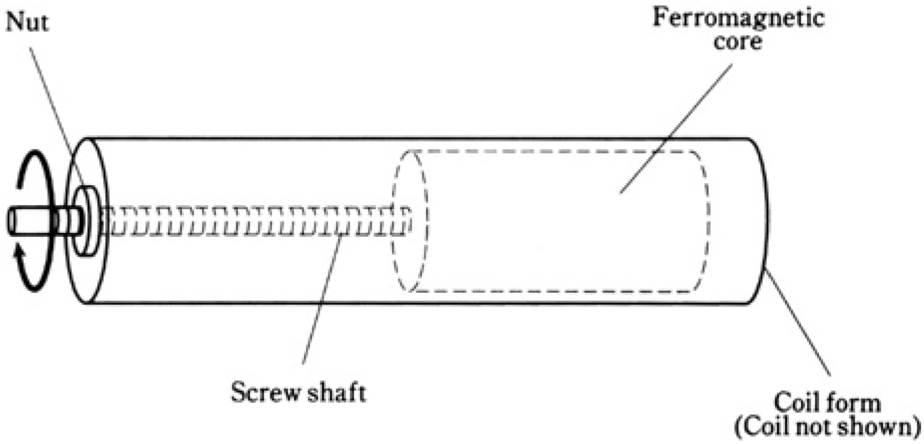
Permeability tuning

Solenoidal, or cylindrical, coils can be made to have variable inductance by sliding ferromagnetic cores in and out of them. This is a common practice in radio communications. The frequency of a radio circuit can be adjusted in this way, as you'll learn later in this book.

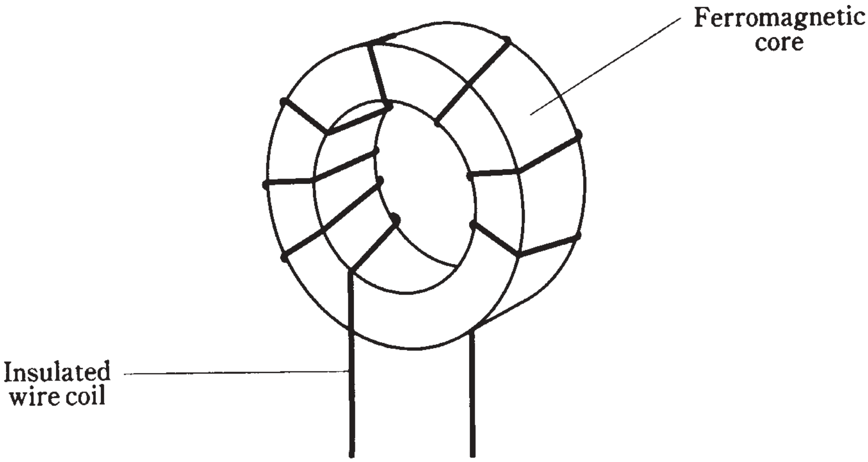
Because moving the core in and out changes the effective permeability within a coil of wire, this method of tuning is called *permeability tuning.* The in/out motion can be precisely controlled by attaching the core to a screw shaft, and anchoring a nut at one end of the coil (Fig. 10-7). As the screw shaft is rotated clockwise, the core enters the coil, so that 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. In recent years, a new form of coil has become increasingly common. This is the *toroid.* It gets its name from the donut shape of the ferromagnetic core. The coil is wound over a core having this shape (Fig. 10-8).



10-7 Permeability tuning of a solenoidal coil.



10-8 A toroidal coil winding.

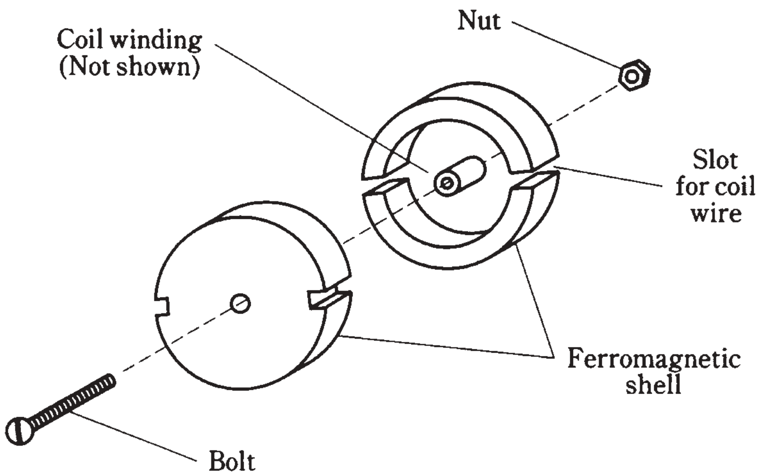
There are several advantages to toroidal coils over solenoidal, or cylindrical, ones. First, fewer turns of wire are needed to get a certain inductance with a toroid, as compared with a solenoid. Second, a toroid can be physically smaller for a given inductance and current-carrying capacity. Third, and perhaps most important, essentially all of the flux in a toroidal inductor is contained within the core material. This reduces unwanted mutual inductances with components near the toroid.

There are some disadvantages, or limitations in the flexibifity, of toroidal coils. It is more difficult to permeability-tune a toroidal coil than it is to tune a solenoidal one. It's been done, but the hardware is cumbersome. Toroidal coils are harder to wind than solenoidal ones.

Sometimes, mutual inductance between or among physically separate coils is wanted; 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. This is to extend a solenoidal core completely around the outside of the coil, making the core into a shell (Fig. 10-9). This is known as a *pot core*. Whereas in most inductors the coil is wound around the form, in a pot core the form is wrapped around the coil.



10-9 A pot core shell. Coil winding is not shown.

The 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.

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 an extremely large inductance within a small volume of 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.

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 audio frequency

Inductors at audio frequencies 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.

Inductors can be used in conjunction with moderately large values of capacitance in order to obtain audio *tuned circuits.* However, in recent years, audio tuning has been taken over by active components, particularly *integrated circuits.*

Inductors at radio frequency

The radio frequencies range from 9 kHz to well above 100 GHz. At the low end of this range, inductors are similar to those at audio frequencies. As the frequency increases, cores having lower permeability are used. Toroids are quite common up through about 30 MHz. Above that frequency, air-core coils are more often used.

In radio-frequency (rf) circuits, 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, while rejecting signals at other frequencies. You'll learn more about this in the chapter on *resonance.*

Transmission-line inductors

At radio frequencies of more than 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.

Types of transmission line

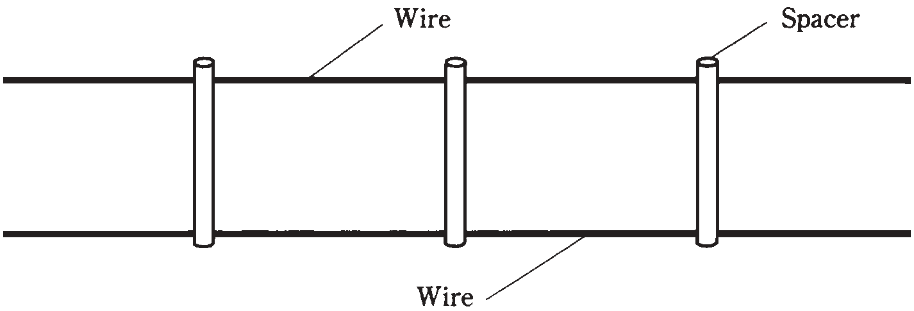
Transmission lines usually take either of two forms, the *parallel-wire* type or the *coaxial* type.

A parallel-wire transmission line consists of two wires running alongside each other with a constant spacing (Fig. 10-10). The spacing is maintained by polyethylene rods molded at regular intervals to the wires, or by a solid web of polyethylene. You have seen this type of line used with television receiving antennas. The substance separating the wires is called the *dielectric* of the transmission line.

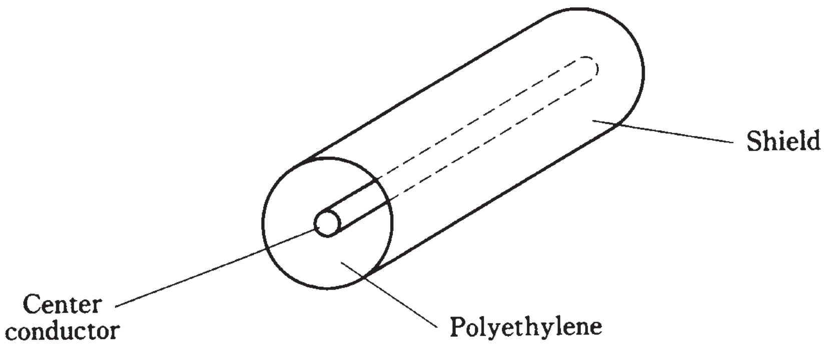
A coaxial transmission line has a wire conductor surrounded by a tubular braid or pipe (Fig. 10-11). The wire is kept at the center of this tubular *shield* by means of polythylene beads, or more often, by solid or foamed polyethylene dielectric, all along the length of the line.

Line inductance

Short lengths of any type of transmission line behave as inductors, as long as the line is less than 90 electrical degrees in length. At 100 MHz, 90 electrical degrees, or ¼



10-10 Parallel-wire transmission line.



10-11 Coaxial transmission line.

wavelength, in free space is just 75 cm, or a little more than 2 ft. In general, if f is the frequency in megahertz, then ¼ wavelength (s) in free space, in centimeters, is given by

$$s = 7500/f$$

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, ¼ wavelength along the line can be anywhere from about 0.66 of the free-space length (for coaxial lines with solid polyethylene dielectric) to about 0.95 of the free-space length (for parallel-wire line with spacers molded at intervals of several inches).

The factor by which the wavelength is shortened is called the velocity factor of the line. This is because the shortening of the wavelength is a result of a slowing-down of the speed with which the radio signals move in the line, as compared with their speed in space (the speed of light). If the velocity factor of a line is given by v, then the above formula for the length of a quarter-wave line, in centimeters, becomes

$$s = 7500v/f$$

Very short lengths of line—a few electrical degrees—produce small values of inductance. As the length approaches ¼ wavelength, the inductance increases.

Transmission line inductors behave differently than coils in one important way: the inductance of a transmission-line section changes as the frequency changes. At first, the inductance will become larger as the frequency increases. At a certain limiting frequency, the inductance becomes infinite. Above that frequency, the line becomes *capacitive* instead. You'll learn about capacitance shortly.

A detailed discussion of frequency, transmission line type and length, and inductance is beyond the level of this book. Texts on radio engineering are recommended for further information on this subject.

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 therefore more significant at radio frequencies than at audio frequencies.

In some cases, especially in radio communications equipment, the inductance of, and among, wires can become a major bugaboo. 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 *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. In some cases, stray inductance can cause life-threatening malfunctions. This might happen with certain medical devices.

The most common 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.

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 high-frequency ac.
   D. Introducing resistance into a circuit.

2. Which of the following does *not* affect the inductance of a coil?
   A. The diameter of the wire.
   B. The number of turns.
   C. The type of core material.
   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. Two inductors are connected in series, without mutual inductance. Their values are 33 mH and 55 mH. The net inductance of the combination is:
   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. Three inductors are connected in series without mutual inductance. Their values are 4 nH, 140 μH, and 5 H. For practical purposes, the net inductance will be very close to:
   A. 4 nH.
   B. 140 μH.
   C. 5 H.
   D. None of these.

9. Suppose the three inductors mentioned above are connected in parallel without mutual inductance. The net inductance will be close to:
   A. 4 nH.
   B. 140 μH.

C. 5 H.

D. None of these.

10. Two inductors, each of 100 μH, are in series. The coefficient of coupling is 0.40. The net inductance, if the coil fields reinforce each other, is:

A. 50 μH.

B. 120 μH.

C. 200 μH.

D. 280 μH.

11. If the coil fields oppose in the foregoing series-connected arrangement, the net inductance is:

A. 50 μH.

B. 120 μH.

C. 200 μH.

D. 280 μH.

12. Two inductors, having values of 44 mH and 88 mH, are connected in series with a coefficient of coupling equal to 1.0 (maximum possible mutual inductance). If their fields reinforce, the net inductance (to two significant digits) is:

A. 7.5 mH.

B. 132 mH.

C. 190 mH.

D. 260 mH.

13. If the fields in the previous situation oppose, the net inductance will be:

A. 7.5 mH.

B. 132 mH.

C. 190 mH.

D. 260 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:

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 winding is:
    A. High current capacity.
    B. Large inductance in small volume.
    C. 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 in a small volume.
    D. Has permeability that can vary over a wide range.

18. At a frequency of 400 Hz, 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 MHz, the best form for an inductor would be:
    A. Air-core.
    B. Pot core.
    C. Either of the above.
    D. Neither of the above.

20. A transmission-line inductor made from coaxial cable, having velocity factor of 0.66, and working at 450 MHz, would be shorter than:
    A. 16.7 m.
    B. 11 m.
    C. 16.7 cm.
    D. 11 cm.

11

CHAPTER

Capacitance

ELECTRICAL COMPONENTS CAN OPPOSE AC IN THREE DIFFERENT WAYS, TWO OF which you've learned about already.

Resistance slows down the rate of transfer of charge carriers (usually electrons) by "brute force." In this process, some of the energy is invariably converted from electrical form to heat. Resistance is said to *consume power* for this reason. Resistance is present in dc as well as in ac circuits, and works the same way for either direct or alternating current.

Inductance impedes the flow of ac charge carriers by temporarily storing the energy as a magnetic field. But this energy is given back later.

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. This energy is given back later, just as it is in an inductance. Capacitance is not generally important in pure-dc circuits. It can have significance in circuits where dc is pulsating, and not steady.

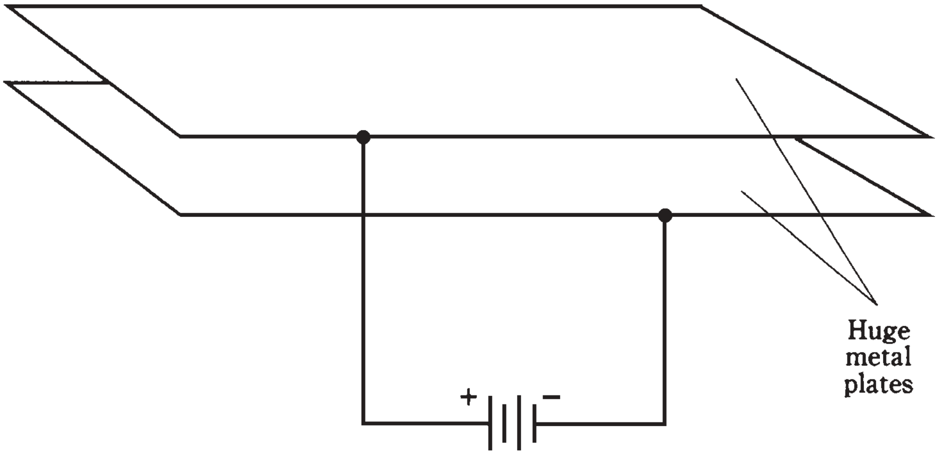
Capacitance, like inductance, can appear when it is not wanted or intended. As with inductance, this effect tends to become more evident as the ac frequency increases.

The property of capacitance

Imagine two very large, flat sheets of metal such as copper or aluminum, 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 just a foot of space. What will happen if these two sheets of metal are connected to the terminals of a battery, as shown in Fig. 11-1?

The two plates will become charged electrically, one positively and the other negatively. You might think that this would take a little while, because the sheets are so big. This is an accurate supposition.

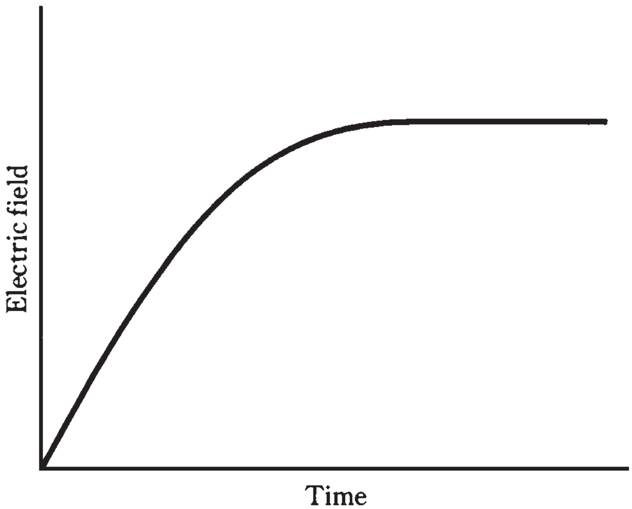
If the plates were small, they would both become charged almost instantly, attaining a relative voltage equal to the voltage of the battery. But because the plates are gigantic,



11-1 A huge pair of parallel plates illustrates the principle of capacitance. See text.

it will take awhile for the negative one to "fill up" with electrons, and it will take an equal amount of time for the other one to get electrons "sucked out." Finally, however, the voltage between the two plates will be equal to the battery voltage, and an electric field will exist in the space between the plates.

This electric field will be small at first; the plates don't charge 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.



11-2 Relative electric field intensity between the huge metal plates, as a function of time.

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*. It is denoted by the letter C.

Practical capacitors

It's out of the question to make a capacitor of the above 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. In fact, two sets of several plates each can be meshed together, with air in between them, and the resulting capacitance will be 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. Plastics work very well for this purpose. 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, it takes some time before the electric field reaches its full intensity. The voltage 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 amount of current flowing and the rate of voltage change across the plates of a capacitor. A capacitance of one *farad,* abbreviated F, represents a current flow of one ampere while there is a potential-difference increase or decrease of one volt per second. A capacitance of one farad also results in one volt of potential difference for an electric charge of one coulomb.

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 a millionth (10^{-6}) of a farad, and 1 pF is a millionth of a microfarad, or a trillionth of a farad (10^{-12} F).

Some quite large capacitances can be stuffed into physically small components. Conversely, some capacitors with small values take up large volumes. The bulkiness of a capacitor is proportional to the voltage that it can handle, more than it is related to the capacitance. The higher the rated voltage, the bigger, physically, the component will be.

Capacitors in series

With capacitors, there is almost never any mutual interaction. This makes capacitors somewhat easier to work with than inductors.

Capacitors in series add together like resistors in parallel. If you connect two capacitors of the same value in series, the result will be half the capacitance of either component alone. In general, if there are several capacitors in series, the composite value will be less than any of the single components. It is important that you always use the same size units when determining the capacitance of any combination. Don't mix microfarads with picofarads. The answer that you get will be in whichever size units you use for the individual components.

Suppose you have several capacitors with values $C_1, C_2, C_3, \ldots C_n$ all connected in series. Then you can find the reciprocal of the total capacitance, $1/C$, using the following formula:

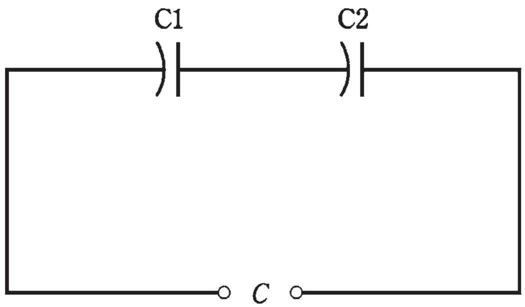
$$1/C = 1/C_1 + 1/C_2 + 1/C_3 + \ldots + 1/C_n$$

The total capacitance, 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 extremely tiny compared with the values of all the others, the composite capacitance can be taken as the value of the smallest component.

Problem 11-1

Two capacitors, with values of $C_1 = 0.10\ \mu\text{F}$ and $C_2 = 0.050\ \mu\text{F}$, are connected in series (Fig. 11-3). What is the total capacitance?



11-3 Capacitors in series.

Using the above 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 $C = 1/30 = 0.033\ \mu\text{F}$.

Problem 11-2

Two capacitors with values of 0.0010 μF and 100 pF are connected in series. What is the total capacitance?

Convert to the same size units. A value of 100 pF represents 0. 000100 μF. Then you can say that $C_1 = 0.0010\ \mu\text{F}$ and $C_2 = 0.00010\ \mu\text{F}$. The reciprocals are $1/C_1 = 1000$ and

$1/C_2 = 10{,}000$. Therefore, $1/C = 1{,}000 + 10{,}000 = 11{,}000$, and $C = 0.\ 000091\ \mu F$. This number is a little awkward, and you might rather say it's 91 pF.

In the above problem, you could have chosen pF to work with, rather than μF. In either case, there is some tricky decimal placement involved. It's important to double-check calculations when numbers get like this. Calculators will take care of the decimal placement problem, sometimes using exponent notation and sometimes not, but a calculator can only work with what you put into it! If you put a wrong number in, you will get a wrong answer, perhaps off by a factor of 10, 100, or even 1,000.

Problem 11-3

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 total value C is just $1/n$ of the capacitance of any of the components alone. Because there are five 100-pF capacitors here, the total is $C = 100/5 = 20$ pF.

Capacitors in parallel

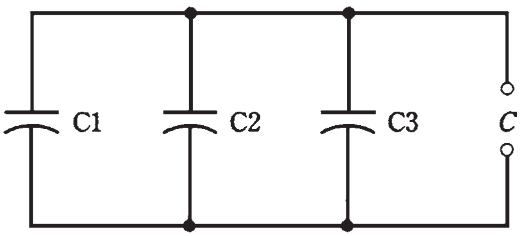
Capacitances in parallel add like resistances in series. That is, the total capacitance is the sum of the individual component values. Again, you need to be sure that you use the same size units all the way through.

If two or more capacitors are connected in parallel, and one of the components is much, much larger than any of the others, the total capacitance can be taken as simply the value of the biggest one.

Problem 11-4

Three capacitors are in parallel, having values of $C_1 = 0.100\ \mu F$, $C_2, = 0.0100\ \mu F$, and C_3 0.00100 μF, as shown in Fig. 11-4. What is the total capacitance?

11-4 Capacitors in parallel.



Just add them up: $C = 0.100 + 0.0100 = 0.00100 = 0.\ 111000$. Because the values are given to three significant figures, the final answer should be stated as $C = 0.111\ \mu F$.

Problem 11-5

Two capacitors are in parallel, one with a value of 100 μF and one with a value of 100 pF. What is the effective total capacitance?

In this case, without even doing any calculations, you can say that the total is 100 μF for practical purposes. The 100-pF unit is only a millionth of the capacitance of the 100-μF component; therefore, the smaller capacitor contributes essentially nothing to the composite total.

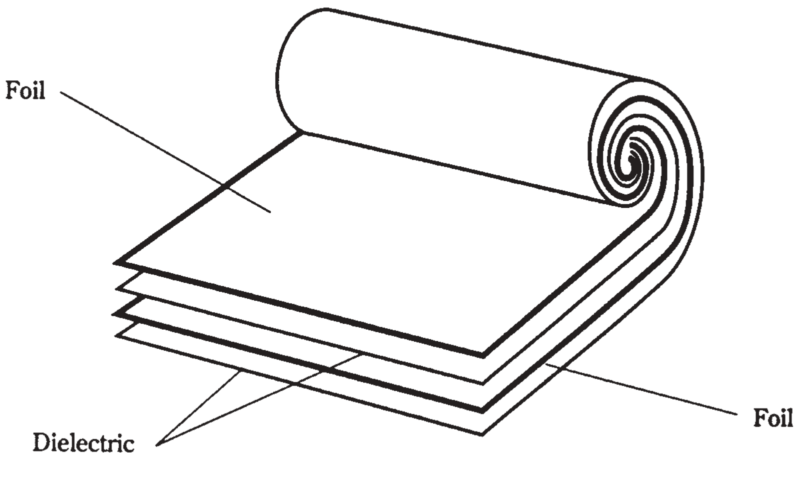
Dielectric materials

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 works quite well as a dielectric. It has almost no loss. But it is difficult to get very much capacitance using air as the dielectric. Some solid material is usually employed as the dielectric for most *fixed* capacitors, that is, for types manufactured to have a constant, unchangeable value of capacitance.

Dielectric materials conduct electric fields well, but they are not good conductors of electric currents. In fact, the 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). Both of these act to increase the capacitance per unit volume, allowing reasonable capacitances to exist in a small volume.



11-5 Foil sheets can be rolled up with dielectric material sandwiched in between.

Paper capacitors

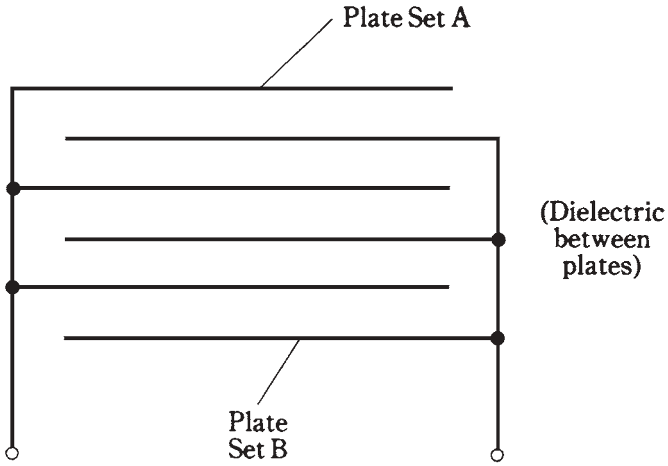
In the early days of radio, capacitors were commonly made by placing paper, soaked with mineral oil, between two strips of foil, rolling the assembly up, attaching wire leads to the two pieces of foil, and enclosing the rolled-up foil and paper in a cylindrical case.

These capacitors can still sometimes be found in 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.

Mica capacitors

When you were a child, you might have seen mica, a naturally occurring, transparent substance that flakes off in thin sheets. This material makes an excellent dielectric for capacitors.

Mica capacitors can be made by alternately stacking metal sheets and layers of mica, or by applying silver ink to the 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.



11-6 Meshing of plates to increase capacitance.

Mica capacitors have low loss; that is, they waste very little power as heat, 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 tend to be 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

Porcelain is another material that works well as a dielectric. 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 quite low loss, and therefore allows for high efficiency.

For low values of capacitance, just one layer of ceramic is needed, and two metal plates can be glued to the disk-shaped porcelain, 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

Various different plastics make good dielectrics for the manufacture of capacitors. Polyester, polyethylene, and polystyrene are commonly used. The substance called *mylar* that you might have seen used to tint windows makes a good dielectric for capacitors.

The method of manufacture is similar to that for paper capacitors when the plastic is flexible. Stacking methods can be used if the plastic is more rigid. The geometries can vary, and these capacitors are therefore found in several different shapes.

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 audio and radio frequencies, 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 above-mentioned 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 considerably greater capacitance than any of the above types, but it must be connected in the proper direction in a circuit to work right. Therefore, 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 units can handle thousands of volts. These capacitors are most often seen in audio-frequency 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 might 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 do not fail often. They can be used in audio-frequency and digital circuits in place of aluminum electrolytics.

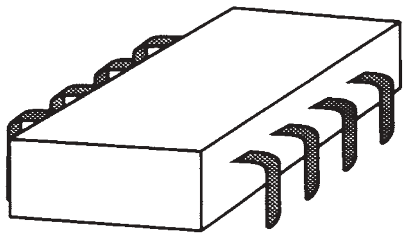
Semiconductor capacitors

A little later in this book, you'll learn about *semiconductors*. These materials, in their many different forms, have revolutionized electrical and electronic circuit design in the past several decades.

These 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 a variable capacitor. Some diodes are especially manufactured to serve this function. Their capacitances fluctuate rapidly along with pulsating dc. These components are called *varactor diodes* or simply *varactors*.

Capacitors can be formed in the semiconductor materials of an *integrated circuit (IC)* 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.

You have probably seen ICs in electronic equipment; almost any personal computer has dozens of them. They look like little boxes with many prongs (Fig. 11-7).



11-7 An integrated-circuit package.

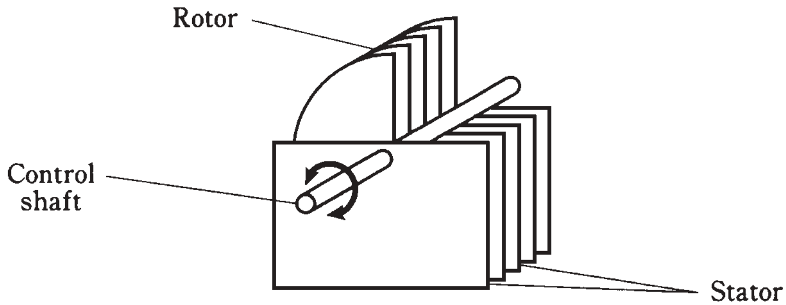
Semiconductor capacitors usually have small values of capacitance. They are physically tiny, and can handle only low voltages. The advantages are miniaturization, and an ability, in the case of the varactor, to change in value at a rapid rate.

Variable capacitors

Capacitors can be varied in value 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 might also 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 rendition of an air-variable capacitor.



11-8 Simplified 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 pF to about 1,000 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 at radio frequencies. They are highly efficient, and are nonpolarized, although the rotor is usually connected to common ground (the chassis or circuit board).

Trimmers

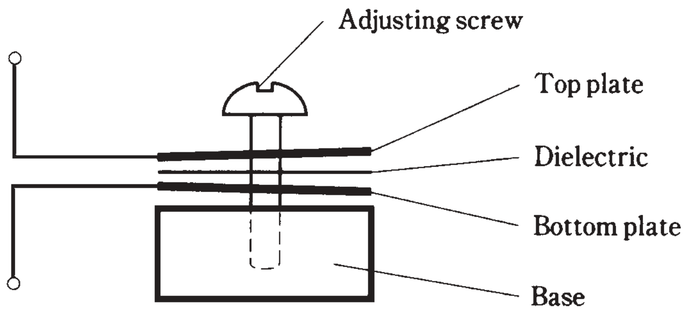
When it is not necessary to change the value of a capacitor very often, a trimmer might be used. It consists of two plates, mounted on a ceramic base and separated by a sheet of mylar, mica, or some other dielectric. The plates are "springy" and can be squashed together more or less by means of a screw (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, and are highly efficient. They are nonpolarized.

Coaxial capacitors

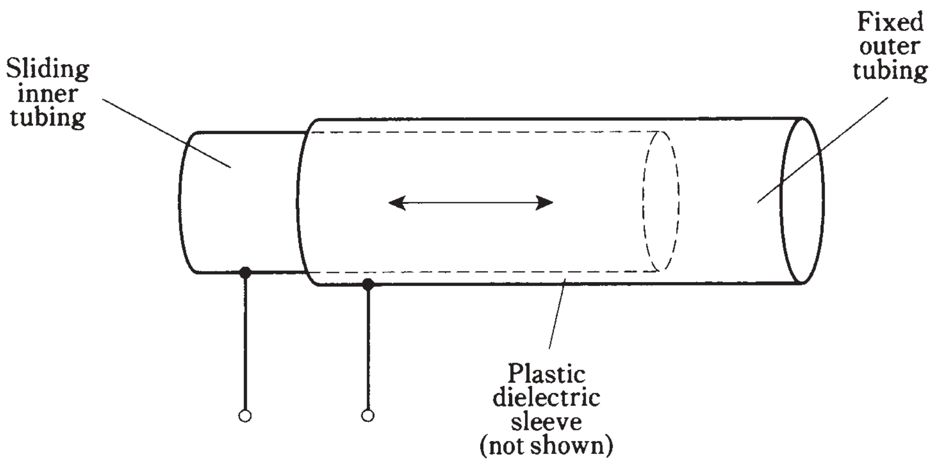
Recall from the previous chapter that sections of transmission lines can work as inductors. They can act as capacitors too.



11-9 A trimmer capacitor.

If a section of transmission line is less than 1/4 wavelength long, and is left open at the far end (rather than shorted out), it will act as a capacitor. The capacitance will increase with length.

The most common transmission-line capacitor uses two telescoping sections of tubing. This is called a *coaxial capacitor* and 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 coaxial variable capacitor.

Coaxial capacitors are used in radio-frequency applications, particularly in antenna systems. Their values are generally from a few picofarads up to about 100 pF.

Tolerance

Capacitors are rated according to how nearly their values can be expected to match the rated capacitance. The most common tolerance is 10 percent; some capacitors are rated at 5 percent or even at 1 percent. In all cases, the tolerance ratings are *plus-or-minus*.

Therefore, a 10-percent capacitor can range from 10 percent less than its assigned value to 10 percent more.

Problem 11-6

A capacitor is rated at 0.001 μF, plus-or-minus 10 percent. What is the actual range of capacitances it can have?

First, multiply 0.001 by 10 percent to get the plus-or-minus variation. This is $0.001 \times 0.1 = 0.0001$ μF. Then add and subtract this from the rated value to get the maximum and minimum possible capacitances. The result is 0.0011 μF to 0.0009 μF.

You might prefer to work with picofarads instead of microfarads, if the small numbers make you feel uneasy. Just change 0.001 μF to 1000 pF. Then the variation is plus-or-minus $1000 \times 0.1 = 100$ pF, and the range becomes 1100 pF to 900 pF.

Temperature coefficient

Some capacitors increase in value as the temperature increases. These components have a *positive temperature coefficient.* Some capacitors' values get less as the temperature rises; these have *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.

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 will cancel each other 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.

You won't have to do calculations involving temperature coefficients if you're in a management position; you can delegate these things to the engineers. If you plan to become an engineer, you'll most likely have computer software that will perform the calculations for you.

Interelectrode capacitance

Any two pieces of conducting material, when they are brought near each other, will 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 ac circuits and at audio frequencies, interelectrode capacitance is not usually significant. But it can cause problems at radio frequencies. The chances for trouble increase as the frequency increases. The most common phenomena are *feedback,* and a change in the frequency characteristics of a circuit.

Interelectrode capacitance is minimized by keeping wire leads as short as possible. It can also be reduced by using shielded cables and by enclosing circuits in metal housings if interaction might produce trouble. This is why, if you've ever opened up a sophisticated communications radio, you might have seen numerous metal enclosures inside the main box.

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 voltage-handling ability increases.
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 voltage-handling ability 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:
   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:
   A. 33 pF.

B. 330 pF.
C. 3300 pF.
D. 33,000 pF.

7. Five 0.050-μF capacitors are connected in parallel. The total capacitance is:
   A. 0.010 μF.
   B. 0.25 μF.
   C. 0.50 μF.
   D. 0.025 μF.
8. If the same five capacitors are connected in series, the total capacitance will be:
   A. 0.010 μF.
   B. 0.25 μF.
   C. 0.50 μF.
   D. 0.025 μF.
9. Two capacitors are in series. Their values are 47 pF and 33 pF. The composite value is:
   A. 80 pF.
   B. 47 pF.
   C. 33 pF.
   D. 19 pF.
10. Two capacitors are in parallel. Their values are 47 pF and 470 μF. The combination capacitance is:
   A. 47 pF.
   B. 517 pF.
   C. 517 μF.
   D. 470 μF.
11. Three capacitors are in parallel. Their values are 0.0200 μF, 0.0500 μF and 0.10000 μF. The total capacitance is:
   A. 0.0125 μF.
   B. 0.170 μF.
   C. 0.1 μF.
   D. 0.125 μF.
12. Air works well as a dielectric mainly because 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. High efficiency.
    B. Small size.
    C. Capability to handle high voltages.
    D. Low loss.
14. A disk ceramic capacitor might have a value of:
    A. 100 pF.
    B. 33 μF.
    C. 470 μF.
    D. 10,000 μF.
15. A paper capacitor might have a value of:
    A. 0.001 pF.
    B. 0.01 μF.
    C. 100 μF.
    D. 3300 μF.
16. An air-variable capacitor might have a range of:
    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:
    A. Its value decreases as the temperature rises.
    B. Its value increases as the temperature rises.
    C. Its value does not change with temperature.
    D. It must be connected with the correct polarity.
19. A capacitor is rated at 33 pF, plus or minus 10 percent. Which of the following capacitances is outside the acceptable range?
    A. 30 pF.
    B. 37 pF.
    C. 35 pF.
    D. 31 pF.

20. A capacitor, rated at 330 pF, shows an actual value of 317 pF. How many percent off is its value?

   A. 0.039.
   B. 3.9.
   C. 0.041.
   D. 4.1.

12
CHAPTER
Phase

AN ALTERNATING CURRENT REPEATS THE SAME WAVE TRACE OVER AND OVER. Each 360-degree cycle is identical to every other. The wave can have any imaginable shape, but as long as the polarity reverses periodically, and as long as every cycle is the same, the wave can be called true ac.

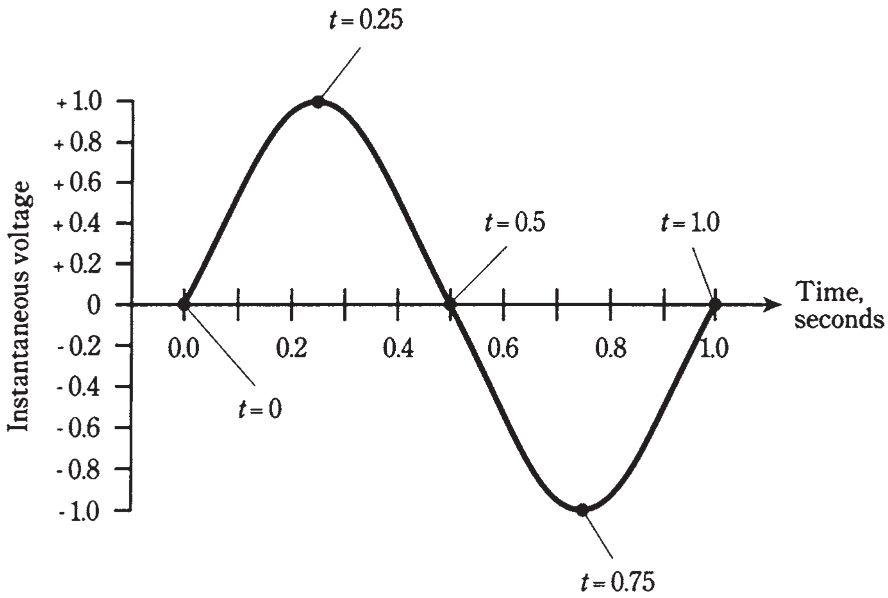
In this chapter, you'll learn more about the most common type of ac, the sine wave. You'll get an in-depth look at the way engineers and technicians think of ac sine waves. There will be a discussion of the circular-motion model of the sine wave. You'll see how these waves add together, and how they can cancel out.

Instantaneous voltage and current

You've seen "stop motion" if you've done much work with a video-cassette recorder (VCR). In fact, you've probably seen it if you've watched any television sportscasts. Suppose that it were possible for you to stop time in real life, any time you wanted. Then you could examine any instant of time in any amount of detail that would satisfy your imagination.

Recall that an ac sine wave has a unique, characteristic shape, as shown in Fig. 12-1. This is the way the graph of the function $y = \sin x$ appears on the coordinate plane. (The abbreviation *sin* stands for *sine* in trigonometry.) Suppose that the peak voltage is plus or minus 1 V, as shown. Further imagine that the period is 1 second, so that 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.

If you freeze time at $t = 446.00$ seconds, say, the voltage will be zero. Looking at the diagram, you can see that the voltage will also be zero every so-many-and-a-*half* seconds; that is, it will be zero at $t = 446.5$ seconds. But instead of getting more positive at these instants, the voltage will be swinging towards the negative.



12-1 A sine wave with period 1 second and frequency 1 Hz.

If you freeze time at so-many-and-a-*quarter* seconds, say $t = 446.25$ seconds, 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 $t = 446.75$ seconds, 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.

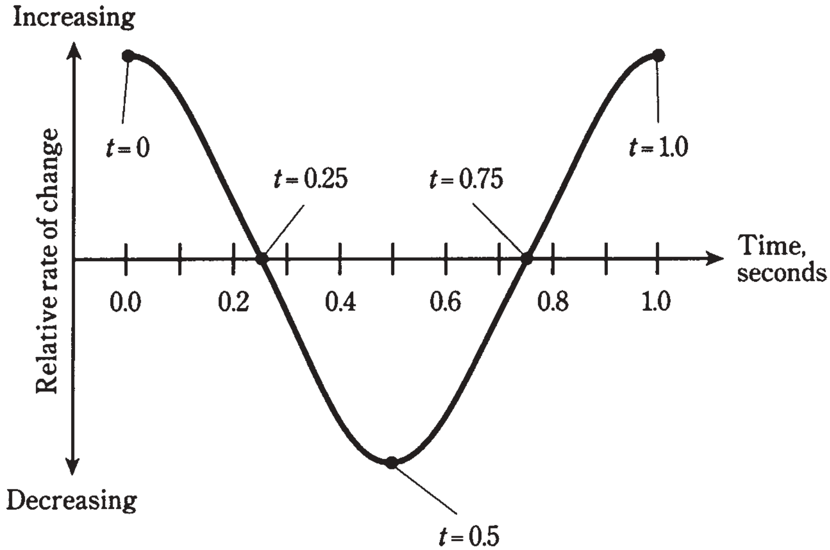
Rate of change

By examining the diagram of Fig. 12-1, you can see 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 $t = 1.0$ in Fig. 12-1. The most rapid decrease takes place when $t = 0.5$.

Notice that when $t = 0.25$, and also when $t = 0.75$, the instantaneous voltage is not changing. This condition exists for a vanishingly small moment. You might liken the value of the voltage at $t = 0.25$ to the altitude of a ball you've tossed straight up into the air, when it reaches its highest point. Similarly, the value of voltage at $t = 0.75$ is akin to the position of a swing at its lowest altitude.

If n is any 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 just 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 plus-or-minus 1 V.

Suppose that you graph the *rate of change in the voltage* of the wave in Fig. 12-1 against time. What will this graph look like? It turns out that it will have a shape that is a sine wave, but it will be displaced to the left of the original wave by one-quarter of a cycle. If you plot the relative rate of change against time as shown in Fig. 12-2, you get the *derivative*, or rate of change, of the sine wave. This is a *cosine* wave, having the same general, characteristic shape as the sine wave. But the *phase* is different.

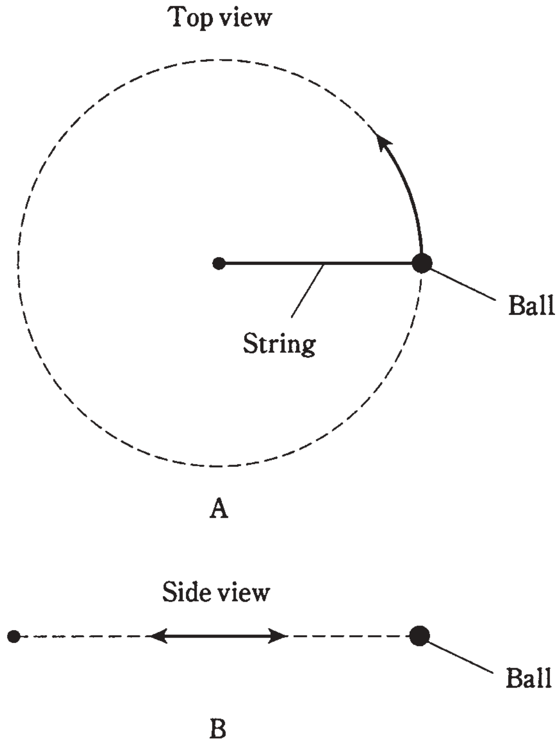


12-2 A sine wave representing the rate of change in instantaneous amplitude of the wave in Fig. 12-1.

Sine waves as circular motion

A sine wave represents the most efficient possible way that a quantity can alternate back and forth. The reasons for this are rather complex, and a thorough discussion of it would get into that fuzzy thought territory where science begins to overlap with esthetics, mathematics, and philosophy. You need not worry about it. You might recall, however, that a sine wave has just one frequency component, and represents a pure wave for this reason.

Suppose that you were to swing a glowing ball around and around at the end of a string, at a rate of one revolution per second. The ball would describe a circle in space (Fig. 12-3A). Imagine that you swing the ball around so that it is always at the same level; that is, so that it takes a path that lies in a horizontal plane. Imagine that you do this in a perfectly dark gymnasium. Now if a friend stands some distance away, with his or her eyes right in the plane of the ball's path, what will your friend see? All that will be visible is the glowing ball, oscillating back and forth (Fig. 12-3B). The ball will seem to move toward the right, slow down, then stop and reverse its direction, going back towards the left. It will move faster and faster, then slower again, reaching its left-most point, at which it will stop and turn around again. This will go on and on, with a frequency of 1 Hz, or a complete cycle per second, because you are swinging the ball around at one revolution per second.



12-3 Swinging ball and string. At A, as seen from above; at B, as seen from edge-on.

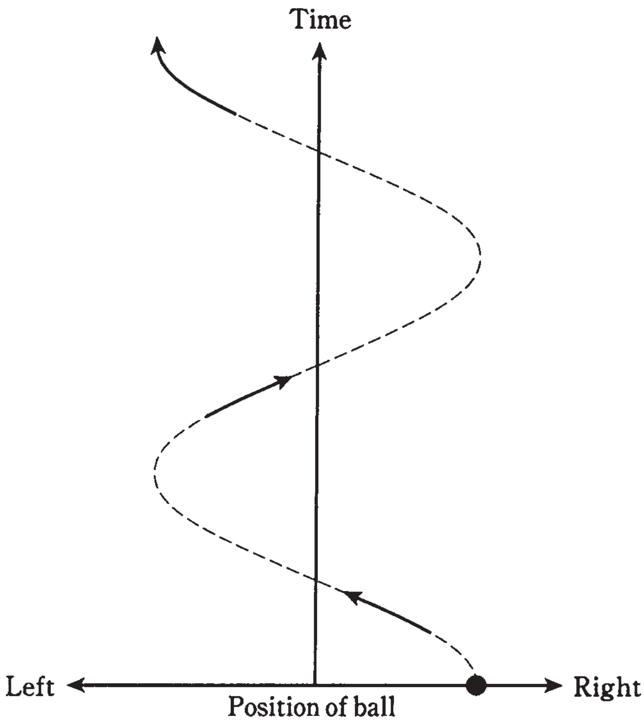
If you graph the position of the ball, as seen by your friend, with respect to time, the result will be a sine wave (Fig. 12-4). This wave has the same characteristic shape as all sine waves.

It is true that some sine waves are taller than others, and some are stretched out horizontally more than others. But the general wave form 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 (x,y) coordinate plane.

You might whirl the ball around faster or slower than one revolution per second. 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 fundamental rule would always apply: the sine wave can be reduced to circular motion.

Degrees of phase

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



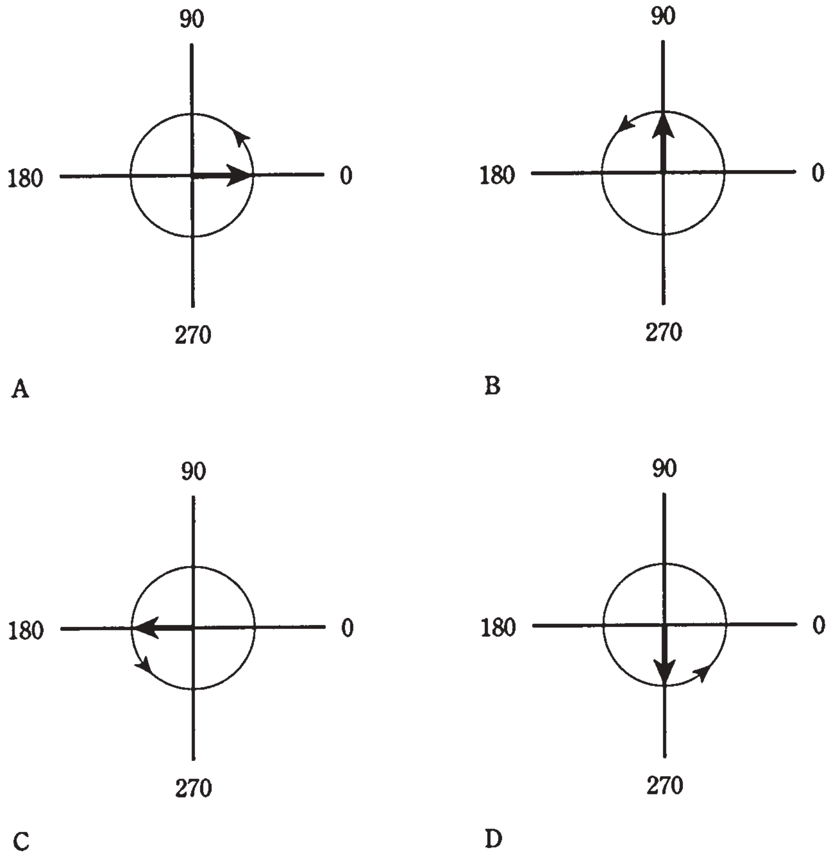
12-4 Position of ball as seen edge-on, as a function of time.

degrees. A sine wave can be represented as circular motion. Exact moments along the sine curve correspond to specific angles, or positions, around a circle.

Figure 12-5 shows the way a *rotating vector* is used to represent a sine wave. At A, the vector points "east," and this is assigned the value of 0 degrees, where the wave amplitude is zero and is increasing positively. At B, the vector points "north"; this is the 90-degree instant, where the wave has attained its maximum positive amplitude. At C, the vector points "west." This is 180 degrees, 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 degrees and represents the maximum negative amplitude. When a full circle has been completed, or 360 degrees, the vector once again points "east." Thus, 360 degrees is the same as 0 degrees. In fact, a value of x degrees represents the same condition as x plus or minus any multiple of 360 degrees.

The four points for the model of Fig. 12-5 are shown on a sine wave graph in Fig. 12-6.

You can think of the vector as going around and around, 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 a revolution per second (1 rps). If the wave has a frequency of 100 Hz, the speed of the vector is 100 rps, or a revolution every 0.01 second. If the wave is 1 MHz, then the speed of the vector is 1,000,000 or 10^6 rps, and it goes once around every 10^{-6}, or 0.000001, second.

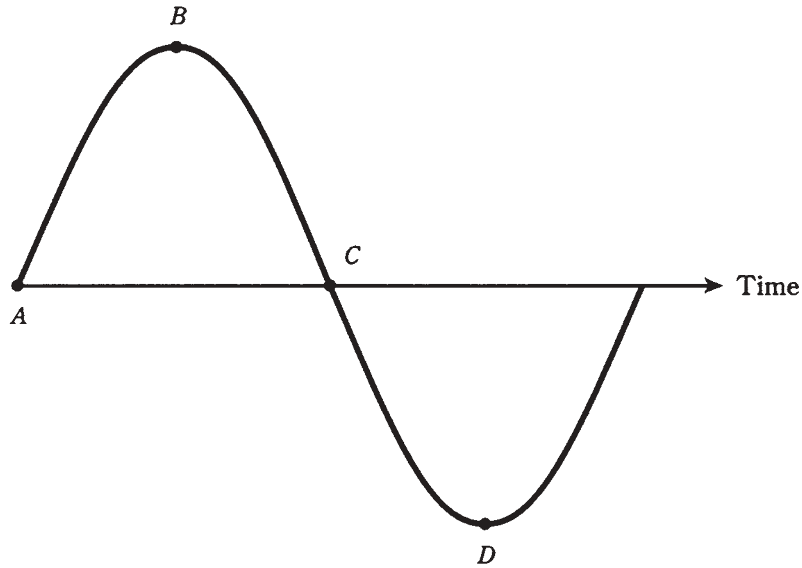


12-5 Vector representation of sine wave at the start of a cycle (A), at ¼ cycle (B), at ½ cycle (C), and at ¾ cycle (D).

The peak amplitude of the wave can be thought of in terms of the length of the vector. In Fig. 12-5, time is represented by the angle counterclockwise from "due east," and amplitude is independent of time. This differs from the more common rendition of the sine wave, such as the one in Fig. 12-6.

In a sense, whatever force "causes" the wave is always there, whether there's any instantaneous voltage or not. The wave is created by angular motion (revolution) of this force. This is visually apparent in the rotating-vector model. The reasons for thinking of ac as a *vector quantity*, having *magnitude* and *direction*, will become more clear in the following chapters, as you learn about reactance and impedance.

If a wave has a frequency of f Hz, then the vector makes a complete 360-degree revolution every 1/f seconds. The vector goes through 1 degree of phase every 1/(360f) seconds.



12-6 The four points for the model of Fig. 12-5, shown on a standard amplitude-versus-time graph of a sine wave.

Radians of phase

An angle of 1 *radian* is about 57.3 degrees. A complete circle is 6.28 radians around. If a wave has a frequency of f Hz, then the vector goes through 1 radian of phase every 1/(57.3f) seconds. The number of radians per second for an ac wave is called the *angular frequency.*

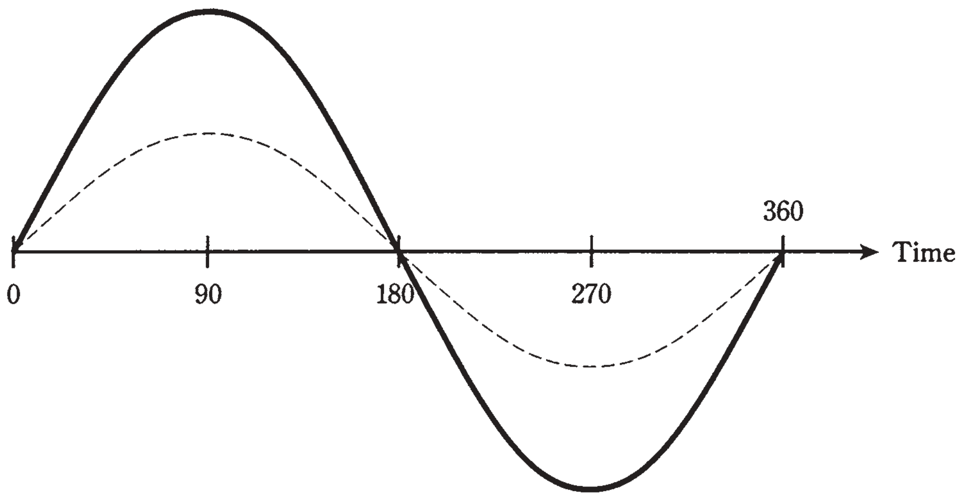
Radians are used mainly by physicists. Engineers and technicians generally use degrees when talking about phase, and Hertz when talking about frequency.

Phase coincidence

When two sine waves have the same frequency, they can behave much differently if their cycles begin at different times. Whether or not the *phase difference,* often called the *phase angle* and specified in degrees, matters depends on the nature of the circuit.

Phase angle can have meaning only when two waves have identical frequencies. If the frequencies differ, even by just a little bit, the relative phase constantly changes, and you can't specify a single number. In the following discussions of phase angle, assume that the two waves always have identical frequencies.

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. (If the amplitudes were the same, you would see only one wave.) The phase difference in this case is 0 degrees. You might say it's any multiple of 360 degrees, too, but engineers and technicians almost never speak of any phase angle of less than 0 or more than 360 degrees.



12-7 Two sine waves in phase coincidence.

If two sine waves are in phase coincidence, the peak amplitude of the resultant wave, which will also be a sine wave, is equal to the sum of the peak amplitudes of the two composite waves. The phase of the resultant is the same as that of the composite waves.

Phase opposition

When two waves begin exactly ½ cycle, or 180 degrees, apart, they are said to be in *phase opposition*. This is illustrated by the drawing of Fig. 12-8. In this situation, engineers sometimes also 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 degrees.

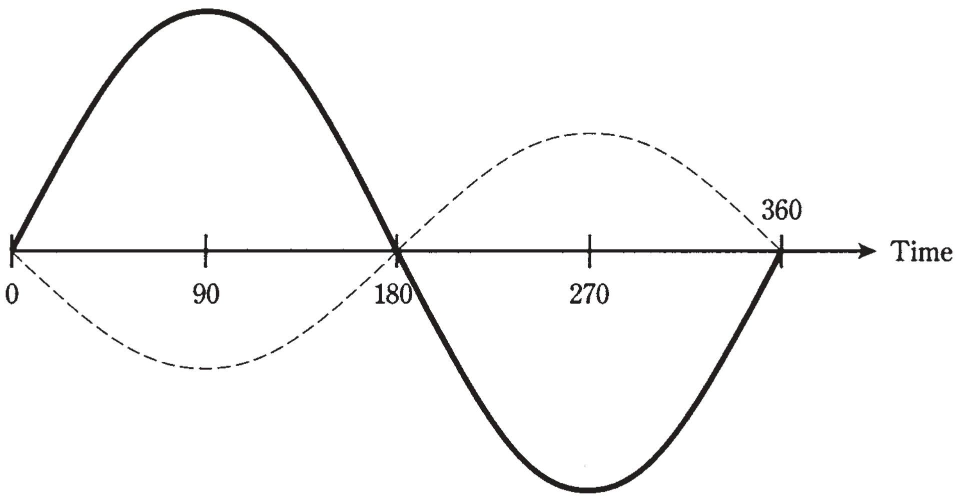
If two sine waves have the same amplitude and are in phase opposition, they will exactly cancel each other out. This is because the instantaneous amplitudes of the two waves are equal and opposite at every moment in time.

If two sine waves have different amplitudes and are in phase opposition, the peak value of the resultant, which will be a sine wave, is equal to the difference between the peak values of the two composite waves. The phase of the resultant will be the same as the phase of the stronger of the two composite waves.

The sine wave has the unique property that, if its phase is shifted by 180 degrees, 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.

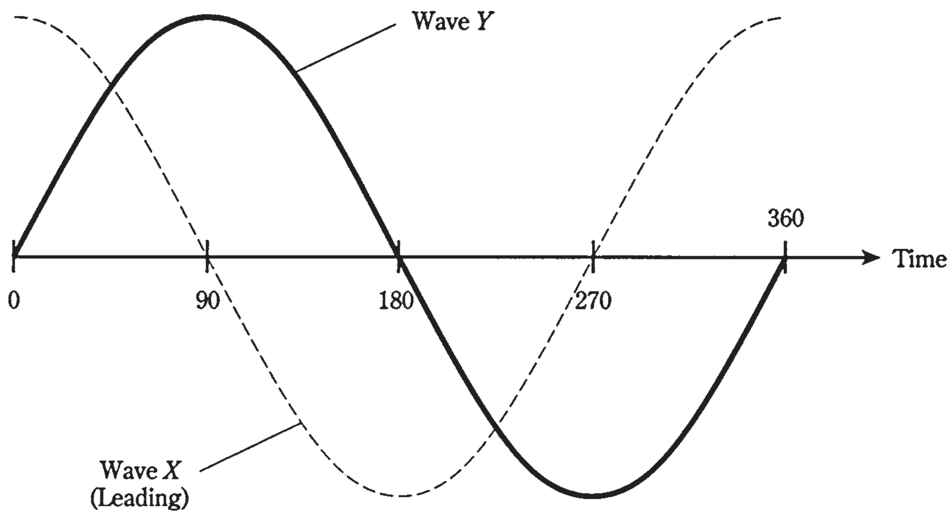
Leading phase

Two waves can differ in phase by any amount from 0 degrees (in phase), through 180 degrees (phase opposition), to 360 degrees (back in phase again).



12-8 Two sine waves in phase opposition.

Suppose there are two sine waves, 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 degrees before Y. Figure 12-9 shows wave X leading wave Y by 90 degrees of phase. The difference could be anything greater than 0 degrees, up to 180 degrees.

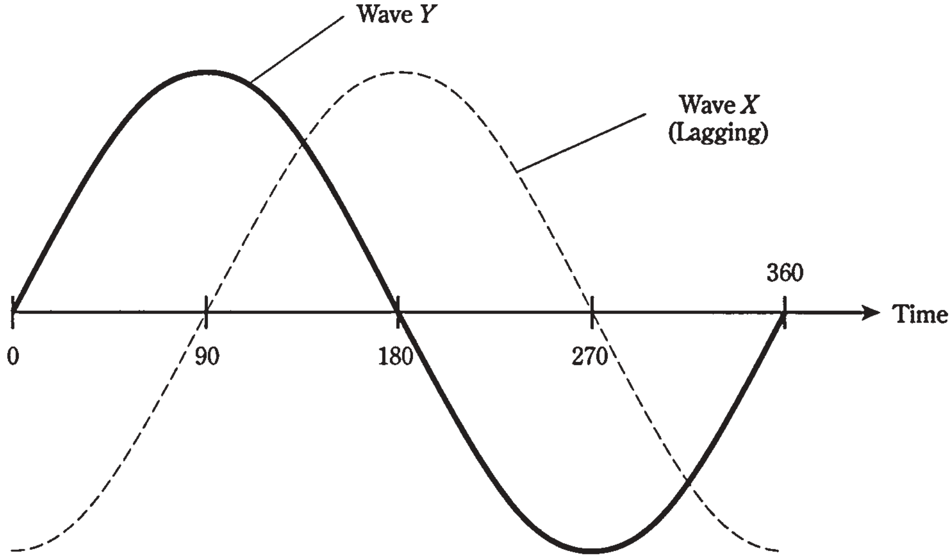


12-9 Wave X leads wave Y by 90 degrees.

Note that if wave X (the dotted line in Fig. 12-9) is leading wave Y(the solid line), then wave X is somewhat to the *left* of wave Y. In a time line, the left is earlier and the right is later.

Lagging phase

Suppose that wave *X* begins its cycle more than 180 degrees, but less than 360 degrees, 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 degrees. Then wave *X* is not leading, but instead is *lagging*, wave *Y*. Figure 12-10 shows wave *X* lagging wave *Y* by 90 degrees. The difference could be anything between 0 and 180 degrees.



12-10 Wave *X* lags wave *Y* by 90 degrees.

You can surmise by now that leading phase and lagging phase are different ways of looking at similar "animals." In practice, ac sine waves are oscillating rapidly, sometimes thousands, millions, or even billions of times per second. If two waves have the same frequency and different phase, how do you know that one wave is really leading the other by some small part of a cycle, instead of lagging by a cycle and a fraction, or by a few hundred, thousand, million, or billion cycles and a fraction? The answer lies in the real-life effects of the waves. Engineers and technicians think of phase differences, for sine waves having the same frequency, as always being between 0 and 180 degrees, either leading or lagging. It rarely matters, in practice, whether one wave started a few seconds earlier or later than the other.

So, while you might think that the diagram of Fig. 12-9 shows wave *X* lagging wave *Y* by 270 degrees, or that the diagram of Fig. 12-10 shows wave *X* leading wave *Y* by 270 degrees, you would get an odd look from an engineer if you said so aloud. And if you said something like "This wave is leading that one by 630 degrees," you might actually be laughed at.

Note that if wave *X* (the dotted line in Fig. 12-10) is lagging wave *Y* (the solid line), then wave *X* is somewhat to the *right* of wave *Y*.

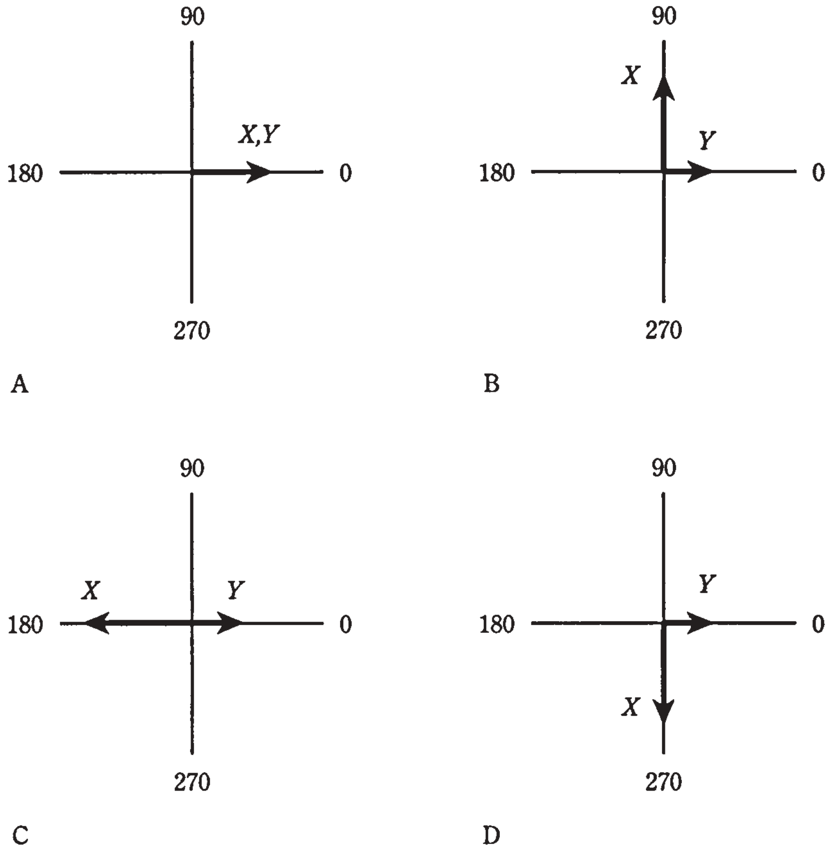
Vector diagrams of phase relationships

The circular renditions of sine waves, such as are shown in the four drawings of Fig. 12-5, are well suited to showing phase relationships.

If a sine wave X is *leading* 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 degrees *counterclockwise* from vector **Y**. If wave X *lags* Y by some number of degrees, then **X** will be *clockwise* from **Y** by that amount.

If two waves are in phase, their vectors overlap (line up). If they are in phase opposition, they point in exactly opposite directions.

The drawings of Fig. 12-11 show four phase relationships between waves X and Y. At A, X is in phase with Y. At B, X leads Y by 90 degrees. At C, X and Y are 180 degrees opposite in phase; at D, X lags Y by 90 degrees. In all cases, you can think of the vectors rotating *counterclockwise* at the rate of f revolutions per second, if their frequency is f Hz.



12-11 Vector representation of phase. At A, waves X and Y are in phase; at B, X leads Y by 90 degrees; at C, X and Y are 180 degrees out of phase; at D, X lags Y by 90 degrees.

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. A sine wave:
   A. Always has the same general appearance.
   B. Has instantaneous rise and fall times.
   C. Is in the same phase as a cosine wave.
   D. Rises very fast, but decays slowly.

3. The derivative of a sine wave:
   A. Is shifted in phase by ½ cycle from the sine wave.
   B. Is a representation of the rate of change.
   C. Has instantaneous rise and fall times.
   D. Rises very fast, but decays slowly.

4. A phase difference of 180 degrees in the circular model represents:
   A. 1/4 revolution.
   B. 1/2 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.
   B. 180.
   C. 270.
   D. 360.

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.
   B. 180.
   C. 270.
   D. 360.

7. A wave has a frequency of 300 kHz. One complete cycle takes:
   A. 1/300 second.
   B. 0.00333 second.
   C. 1/3,000 second.
   D. 0.00000333 second.

8. If a wave has a frequency of 440 Hz, how long does it take for 10 degrees of phase?
   A. 0.00273 second.
   B. 0.000273 second.
   C. 0.0000631 second.
   D. 0.00000631 second.

9. Two waves are in phase coincidence. One has a peak value of 3 V and the other a peak value of 5 V. The resultant will be:
   A. 8 V peak, in phase with the composites.
   B. 2 V peak, in phase with the composites.
   C. 8 V peak, in phase opposition with respect to the composites.
   D. 2 V peak, in phase opposition with respect to the composites.

10. Shifting the phase of an ac sine wave by 90 degrees 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. A phase difference of 540 degrees would more often be spoken of as:
   A. An offset of more than one cycle.
   B. Phase opposition.
   C. A cycle and a half.
   D. 1.5 Hz.

12. 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 a peak amplitude of:
   A. 4 V, in phase with the composites.
   B. 4 V, out of phase with the composites.
   C. 4 V, in phase with wave *X*.
   D. 4 V, in phase with wave *Y*.

13. If wave *X* leads wave *Y* by 45 degrees of phase, 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. *Y* is 120 degrees earlier than *X*.

   B. *Y* is 90 degrees earlier than *X*.

   C. *Y* is 60 degrees earlier than *X*.

   D. *Y* is 30 degrees earlier than *X*.

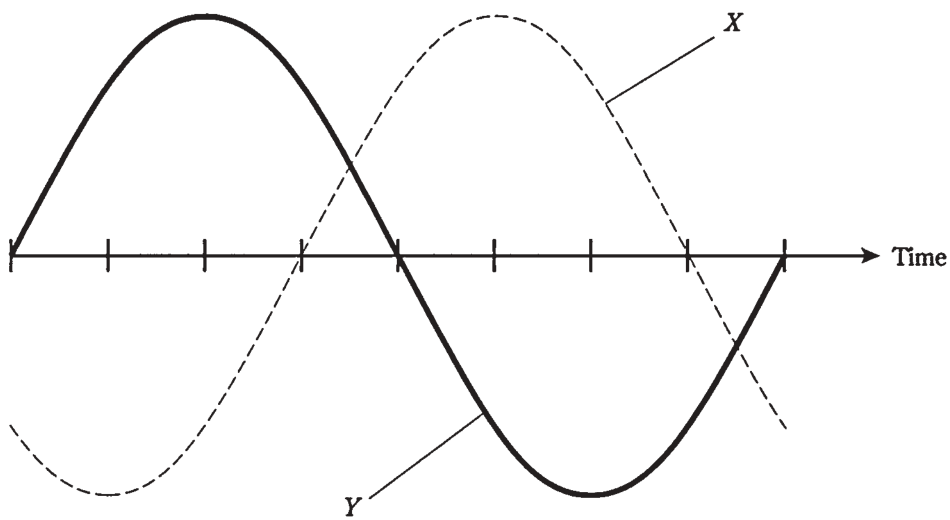
15. In the drawing of Fig. 12-12:

   A. *X* lags *Y* by 45 degrees.

   B. *X* leads *Y* by 45 degrees.

   C. *X* lags *Y* by 135 degrees.

   D. *X* leads *Y* by 135 degrees.



12-12 Illustration for quiz question 15.

16. Which of the drawings in Fig. 12-13 represents the situation of Fig. 12-12?

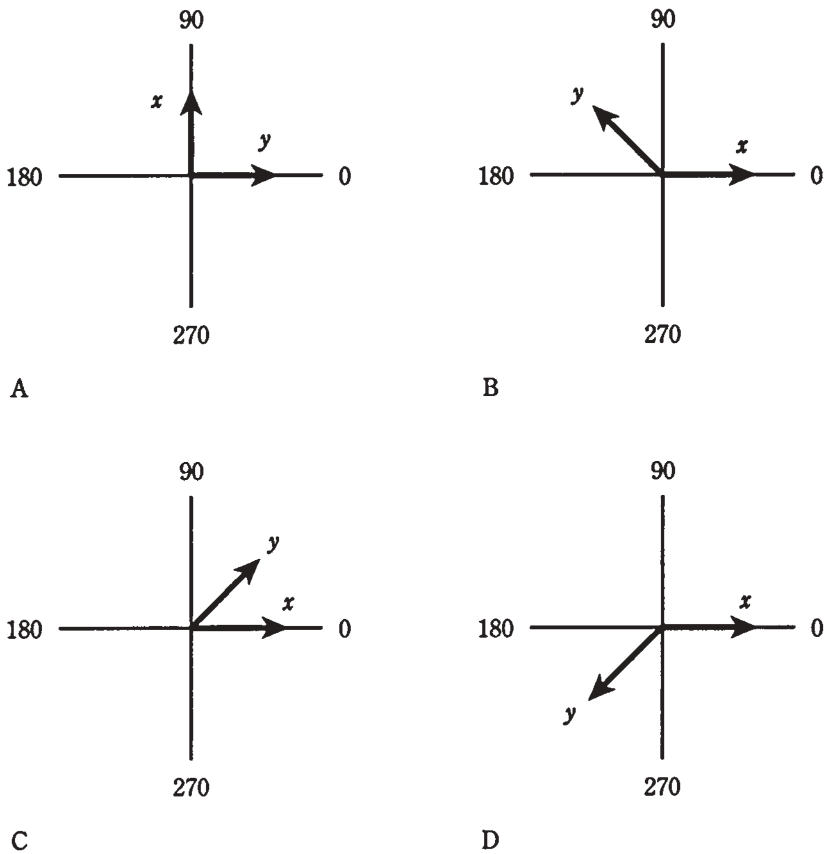
   A. A.

   B. B.

   C. C.

   D. D.

17. In vector diagrams such as those of Fig. 12-13, length of the vector represents:
    A. Average amplitude.
    B. Frequency.
    C. Phase difference.
    D. Peak amplitude.



12-13 Illustration for quiz questions 16 through 20.

18. In vector diagrams such as those of Fig. 12-13, the angle between two vectors represents:
    A. Average amplitude.
    B. Frequency.
    C. Phase difference.
    D. Peak amplitude.