Lesson 13 – Applications of Time-varying Circuits

13.0 Introduction

In this lesson we'll look at a number of applications of time-varying electric and magnetic fields to everyday life. Of course, we can only consider a small number of these applications and not treat any of them in great depth. You will see, however, how the concepts we have learned this semester came together to help understand several these applications.

13.1 Transformers

We found in Lesson 6 that when you put a slab of dielectric material inside a capacitor, it increases its capacitance. We can do something similar with an inductor by putting a soft iron core inside the inductor. Soft iron has a very small coercive force, so it becomes magnetized readily when it is placed in the magnetic field of the inductor. When the domains of the iron are aligned in the inductor's field, they can increase the strength of the magnetic field significantly. When the magnetic flux through the inductor is larger, the induced EMF is larger, and the inductance is larger. Furthermore, the iron core of an inductor tends to keep the magnetic field lines "trapped" inside the iron. What I mean by this can be seen in Fig. 13.1.

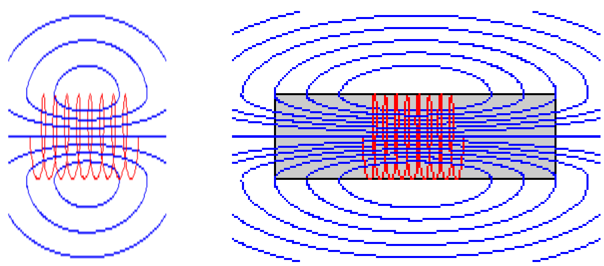


Figure 13.1. The B field of an inductor with and without an iron core.

If we want to keep the magnetic field of the inductor confined in space, we can even make the iron core into a loop. If we do this, the magnetic field lines will follow around the iron loop, as shown in Fig. 13.2.

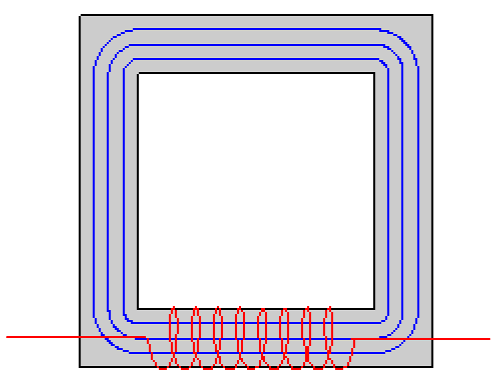


Figure 13.2. An inductor with a looped iron core.

In an idealized inductor of this type, we can determine the magnetic field by applying Ampère's Law in much the same way that we did for a torus. Let's assume that the magnetic field is uniform through the core. The effect of the iron core is to "amplify" the magnetic field in much the same way that the effect of a dielectric in a capacitor is to reduce the electric field. We then have:

$$\Lambda_B = k\mu_0 i_{enc}$$
$$B\ell = k\mu_0 N_1 i$$

where

ℓ is the (average) length of the path around one field line.
N_1 is the total number of turns in the coil of wire.
i is the current passing through the coil.
k is the factor by which the magnetic field is multiplied due to the presence of the ferromagnetic core.

We can also use Faraday's Law to calculate the inductance of this inductor just as we did with an air-core solenoid in Lesson 12. The induced EMF is:

$$V_1 = -N_1 \frac{d\Phi_B}{dt} = -N_1 A \frac{k\mu_0 N_1}{\ell}\frac{di}{dt} = -L\frac{di}{dt}$$
$$\Rightarrow L = \frac{k\mu_0 N_1^2 A}{\ell}$$

Now, we'll take this same inductor and add a second coil around the opposite side of the iron loop and then attach the bottom coil to a power supply. The top coil now has a magnetic field passing through it and this magnetic field constantly changes in time. This, of course, induces a voltage across the upper coil.

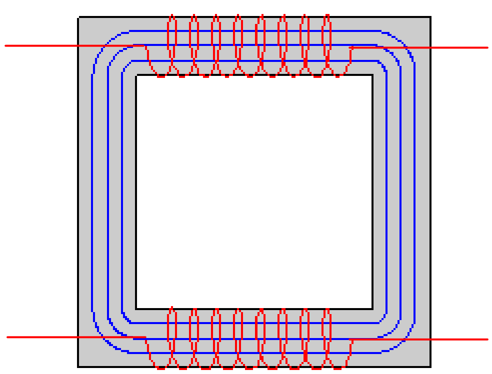


Figure 13.3. An iron-core transformer.

Faraday's Law gives us the EMF across the upper coil.

$$V_2 = -N_2 \frac{d\Phi_B}{dt}$$
$$= N_2 \frac{V_1}{N_1} \quad \text{since} \quad V_1 = -N_1 \frac{d\Phi_B}{dt}$$

(13.1 Voltage in a transformer) $$\frac{V_2}{V_1} = \frac{N_2}{N_1}$$

This relation tells us that by adjusting the ratio of turns in the top coil to the bottom coil, we can get any voltage we want in the top coil. Such a device is called a transformer. Transformers find uses in many different applications. It is helpful to learn a few terms relating to transformers:

• primary coil: the coil that is attached to the power source.
• secondary coil: the coil that is attached to a load.
• step-up transformer: a transformer in which the voltage of the secondary is greater than the voltage in the primary.
• step-down transformer: a transformer in which the voltage of the secondary is less than the voltage in the primary.

The current in the primary and secondary coils of transformers is a little harder to calculate; however, the power provided by the primary is the power available for a load in the secondary. If we assume that the power factors are nearly equal to one for both primary and secondary, this condition becomes:

$$i_1 V_1 = i_2 V_2 .$$

The assumption that the power factors are approximately equal to one is a rather poor assumption, but the relation does tell us something qualitative about the current available in the secondary coil. The current available for the load in step-down transformers is greater in the current in the primary. Conversely, the current available for the load in step-up transformers is less than the current in the primary.

Step-down transformers are used in applications where a higher voltage is not very desirable, such as in a doorbell circuit is a home, or in an adapter for electronic devices. (Such adapters also usually convert AC power to DC power.) Step-down transformers are also used for application like arc welding where large currents are required.

Step-up transformers are used when high voltages are needed in applications, such as in neon signs.

Things to remember:
• Transformers are devices that change AC voltage by taking advantage of mutual induction between two coils wound around an iron core.
• The voltages in the primary and secondary coils are related by the formula $\frac{V_2}{V_1} = \frac{N_2}{N_1}$.
• The power in the primary and secondary coils is approximately related by the formula $i_1 V_1 = i_2 V_2$.

13.2. Getting Electric Power to Your Home

One of the most important applications of electricity and magnetism in everyday life is the electric power we use in our homes and buildings. We will learn most of the science you will need to know to wire a house. Much of what you will learn is based upon simple resistive circuits. One difference between the circuits we studied earlier and the circuits in our homes is that public utilities use alternating current (AC) rather than direct current (DC) power.

A battery produces direct current. That is, the current in a simple resistive circuit is constant: it has the same magnitude and the same direction at all times. Most electronic devices such as computers, digital clocks, radios, etc., require DC power. This means, of course, that something must be done to convert the AC power that comes from the outlet into DC power to be used in these devices. Some motors, such as those in kitchen most appliances, are built to run on AC power directly. Incandescent lights, heaters, etc., can run on either AC or DC power.

In the United States, the standard for outlets is 110–120 V, 60 Hz, AC power. Because it is a convenient number, we will call line voltage 120 V. If we graph the voltage in our outlets as a function of time, we have something like:

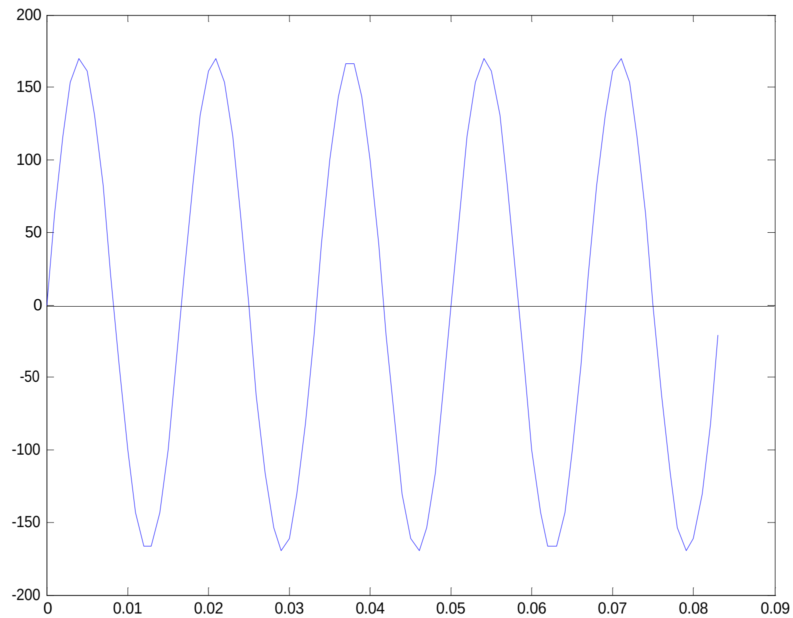


Figure 13.4. AC line voltage as a function of time.

The frequency is f = 60 Hz = 60 cycles/second. The period is the time it takes for the voltage to go through one complete cycle. This is just $T = 1/f$ = 0.167 sec, as can be read (roughly) off the graph. The angular frequency is the number of radians per second, or 2π times the number of cycles per second. In other words, it is $\omega = 2\pi f$ = 377 radians/second. Finally, the maximum voltage is about 170 V. The functional form for an AC voltage is generally of the form:

$$V(t) = 170V\sin\left(377\frac{rad}{s}t\right) = V_{\max}\sin(\omega t)$$

So why do we call this a 120 V outlet? You probably recall from the last chapter that we generally use the rms voltage or current when we deal with AC circuits and $170 \approx 120\sqrt{2}$.

When electric power companies were first begun in the United States, there was a debate over whether DC power or AC power should be provided. Niccolo Tesla advocated AC while Thomas Edison advocated DC power. The resolution of the issue came down to a question of energy loss in power lines. We can consider the generator–transmission line–load system to be a simple series circuit. The load, everyone who is using the power, is a resistor R_l. The transmission line is resistor R_t. The generator is (essentially) a battery with voltage V.

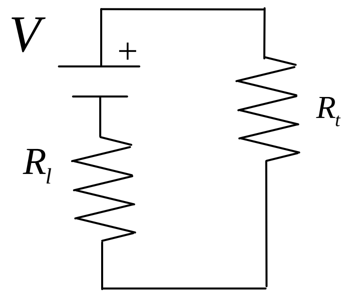


Figure 13.5. A simple circuit we can use to model transmission lines.

Since the algebra can get a little messy, we'll assign some numbers to these quantities. Let V =100 V, and each resistance be 10 Ω. First, we wish to calculate the power consumed by the load and the power lost in the transmission line.

$$I = \frac{V}{R_t + R_l} = 5A$$
$$V_l = IR_l = 50V, \quad V_t = IR_t = 50V$$
$$P_l = IV_l = 250W, \quad P_t = IV_t = 250W, \quad P_b = IV = 500W$$

Now let us increase the voltage of the battery to 250 V while we keep the power used by the load the same. Of course, to keep the power usage the same, we must change the resistance of the load. Leaving all the numbers in SI units (we won't write the units explicitly), we now have:

$$I' = \frac{V'}{R_{tot}} = \frac{250}{10 + R_l'}$$
$$P = 250 = I'^2 R_l' = \left(\frac{250}{10 + R_l'}\right)^2 R_l'$$
$$(10 + R_l')^2 = 250 R_l'$$

The solution of this quadratic equation leads to two possible resistances. Once we know the resistance, we can find the current and the power loss in the transmission line. One value of power loss is very large, so we discard that solution. The second value of resistance gives the following results:

Voltage of the power source	100V	250V
Resistance of the load	10 Ω	229.6 Ω
Power loss in the load	250 W	250 W
Power loss in the line	250 W	10.9 W
Power provided by the battery	500 W	260.9 W

The conclusion is then that higher voltages are best suited for electrical power transmission. On the other hand, if you plug a vacuum cleaner into a 20,000V outlet, you would have to be rather careful. By the time we use electrical power in our homes, we really need to have fairly low voltages. As we from the last section, we can easily change voltages when we use AC power. We simply take a high voltage line into a transformer and out comes whatever voltage we want. High voltage (or "high tension" — tension is the European term for voltage) power lines can use voltages of several hundred kilovolts. Electrical substations have transformers that typically reduce voltages to 4–8 kV. If you look around power lines in residential areas you will see cylinders with power lines going to two or three homes. These cylinders are transformers that take the voltage down to 120 V for use in homes. Actually, three wires go from the transformer to each house. One wire is a ground, 0V. The other two wires have 120 V, but the phases of the voltage are opposite, as shown in Fig. 13.6.

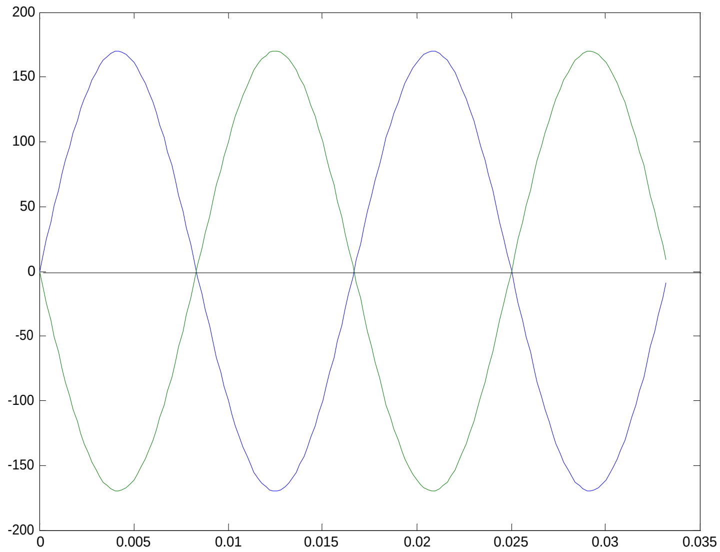


Figure 13.6. Household power with two voltages out of phase by 180º.

If we attach a motor, for example, between these two wires, the voltage across the motor will by 240 V rather than 120 V. This higher voltage is used for major appliances such as electric ovens and electric clothes dryers. (In practice, the voltage is often called 208 V.)

Once the power cables come to a house, they go into an electric meter. This meter measures the amount of energy that is used. We'll learn later how these meters work. After passing through the meter, the power cables enter a service panel in the home. Service panels distribute the power to a number of separate circuits within the home.

Figure 13.7. The route of electric power into a home: high voltage power lines, substation, transformer, service drop, electric meter, service panel (circuit breakers).

Things to remember
• In the United States household power is 120 V (rms), 60 Hz power.
• Transmission lines use high voltages in order to minimize loss of energy in the transmission lines.
• Transformers are used to reduce high voltage power to 120 V for household use.
• Three lines come into a house, 0V and two 120 V that are out of phase by 180 º.

13.3 Circuits and Circuit Breakers

Each circuit that comes from the service panel services a particular part of the home's electrical needs. Large appliances, such as stoves and dryers require their own circuits. Other circuits may include all the lights and outlets in a particular part of a house. When more current flows in a wire, the wire has to be larger in diameter (as we shall soon see). Large wire is more expensive and harder to use because of its size and stiffness. For these reasons, and also just for convenience, houses typically have many different circuits. Circuits for major appliance are usually designed to carry about 40–50 A. Other circuits usually carry no more than 20 A.

On each circuit is a circuit breaker. This is a device that serves two purposes: 1) it is a switch to turn off power to a particular part of a house, and 2) it automatically shuts off power to the circuit if there is too much current flowing through the circuit.

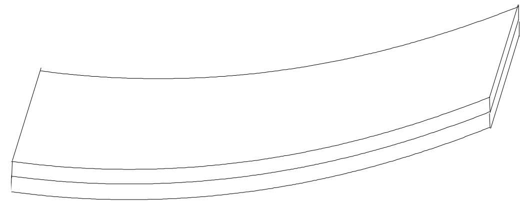


Figure 13.8. A bimetallic strip. The metal on the bottom expands more when it gets hot.

In older homes and in many applications such as in cars, fuses are used instead of circuit breakers. Fuses have strips of metal that melt if current reaches a given value, thereby breaking the circuit. Circuit breakers perform the same function, but they do it in a more sophisticated way that allows them to be simply reset by a switch. Circuit breakers usually have two modes of operation. One mode senses heat and one senses current. If there is enough current to cause a resistor in the breaker to get hot, it causes a bimetallic strip to bend and break the circuit. (A bimetallic strip is a thin strip made of one metal on top and another metal on the bottom. The two metals expand at different rates as the strip heats, causing it to bend. Bimetallic strips are often used in thermostats.) The second mode of operation kicks in when there is a very large current.

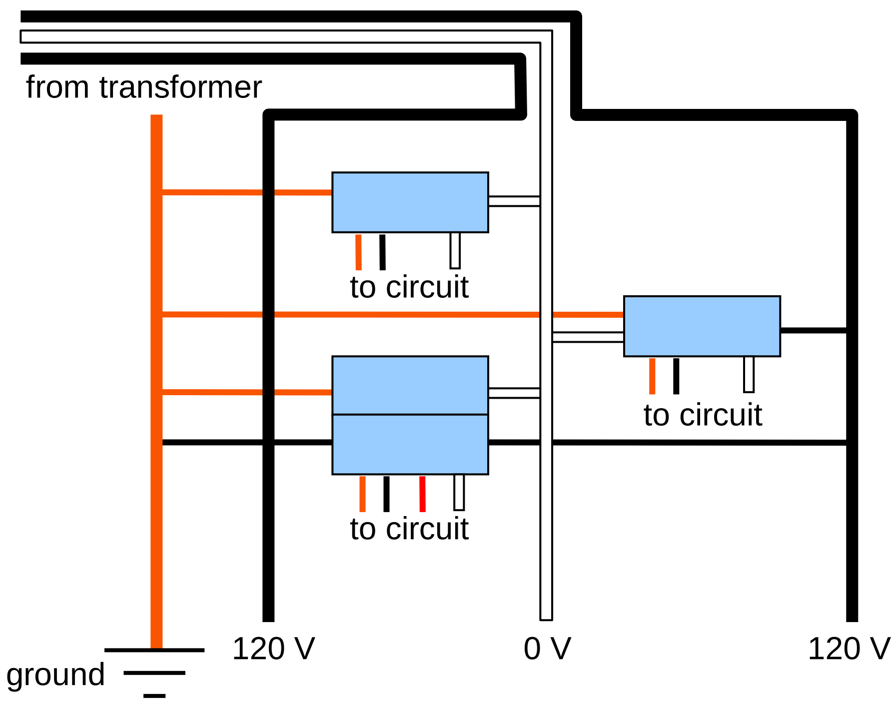


Figure 13.9. Schematic representation of how wires are connected in a service panel. The blue boxes are the circuit breakers. The top circuits are 120 V circuits and the bottom is a 240 V circuit.

In this mode, a small solenoid produces a magnetic field that repels a permanent magnet connected to a switch. When the current gets large, the magnetic field of the solenoid gets large and the switch closes within a fraction of a second. Circuit breakers can be purchased in a variety of sizes, but 15 A and 20 A breakers are used in most home applications.

By appropriate choice of circuit breakers, circuits can be made to operate at either 120 V or 240 V. As we described above, there are three different lines that enter the service panel: one is grounded and the other two each carry 120 V in opposite phases. 120 V circuit breakers connect one 120 V leg to the ground. 240 V circuit breakers connect the two 120 V legs together.

Things to remember:
• Three power lines come into a house. One line is at 0 V and two are at 120 V.
• The 120 V lines are out of phase by 180°. By connecting to both 120 V lines, a 240 V circuit is produced.
• Circuit breakers protect against excessive current in a circuit.
• Home circuit breakers have a bimetallic strip switch and a solenoid switch.

13.4. Wires

As noted in Fig. 13.9, there are three wires that leave the service panel in each 120 V circuit. One wire is connected to the 120 V line. This is called the "hot wire." A second wire is connected to the 0V line from the transformer, so it grounded at the transformer. This is called the "neutral wire." The third wire is connected to a ground in the house and is called the "ground wire." These wires are typically bundled together in a single cable with electrical insulation around the entire bundle. Near the service panel and each outlet, light switch, lamp, etc., the outside insulation is stripped back so the individual wires in the cable can be accessed. Inside the bundle there is one wire with black insulation, one with white insulation, and one that is bare. The wire with black insulation is the hot wire, the one with white insulation is the neutral wire, and the bare wire is the ground.

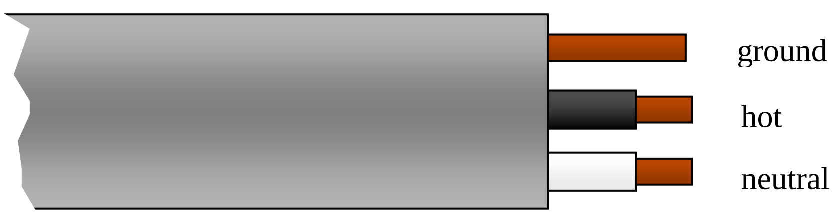


Figure 13.10. Wires bundled in a cable.

Just as a note of caution: when current is flowing through a circuit – through a vacuum cleaner plugged into an outlet, for example – current passes through the hot (black) wire into the vacuum cleaner, and out the neutral (white) wire, and back to the transformer where it passes on to ground. This means that although we call the black wire hot, there can be current passing through the white wire and it can be just as deadly as the black wire! Of course, if you insert a metal object into the neutral side of an outlet, you will be OK, because there is no voltage on the neutral side of the outlet unless there is current in the circuit. But to be on the safe side, you should always treat neutral or white as if it were hot.

In the cable of a 240V circuit, there is an additional wire connected to the second hot wire of the service panel. This fourth wire is colored red to distinguish it from the first wire.

Table 13.1. Color Code for Wires	
black	hot
red	hot, opposite phase
white	neutral
no insulation or green	ground
white with black tape	hot – when both black and white need to be hot.

When we select wire for our circuits, we need to consider two principal things: material and size. The only types of wire that are normally used in wiring are aluminum and copper. Aluminum has the advantage that it is less expensive, but it has several disadvantages: it is more brittle than copper, making it easier to break when you have to bend wires to fit into electrical boxes; aluminum corrodes more easily than copper; aluminum has greater resistivity than copper, so aluminum wires must be of larger diameter than copper; and aluminum wire cannot be safely connected directly to copper wire, as the connection can have high resistance and can overheat. The bottom line is that aluminum wire is often used for the large wire needed in high amperage circuits, but copper is generally used everywhere else. If you do use aluminum wire, however, be sure that all switches, outlets, etc. are designed for use with aluminum.

The next thing to consider is the size of the wire. The concern here is that if we put too much current through a wire, the wire can overheat and cause electrical fires. Since we use circuit breakers in a house, we do know the maximum current that can pass through a wire. The basic physics is fairly simple, but the math is a bit messy. You don't need to worry too much about the details, but try to follow them the best you can. For a fixed amount of current, the power lost as heat in a section of wire is $P = IV = I^2R$ where V is the voltage across the length of wire and R is its resistance. This then represents the amount of heat energy per unit time added to the wire. Heat leaves the length of wire through conduction. The electrical insulation also acts as a thermal insulation that helps hold the heat in the wire. As you may have learned in Physics 123 (Don't worry if you haven't had 123.), the amount of heat per unit time that passes out of the wire is

$$P = kA\frac{T - T_0}{d}$$

where:

k is the thermal conductivity of the insulation
A is the surface are of the length of wire
T is the temperature of the wire, which we take to be uniform
T_0 is the temperature outside the insulation of the wire
d is the thickness of the insulation

Finally, the change in temperature in the wire is related to the net heat flow into the wire by the expression also from Physics 123):

$$Q = mc\Delta T$$

where:

Q is the flowing into the wire
m is the mass of the wire
c is the specific heat of the copper or aluminum
ΔT is the change in temperature of the wire

Let us take the wire to be cylindrical of radius r and length L with a resistivity ρ. Then we have:

$$\frac{Q}{\Delta t} = I^2 R - kA\frac{T - T_0}{d} = mc\frac{\Delta T}{\Delta t}$$
$$mc\frac{\Delta T}{\Delta t} = -\frac{kA}{d}T + \left(I^2 R + \frac{kA}{d}T_0 \right)$$

This is a differential equation that tells us how the temperature of the wire changes in time. It can be solved to yield:

$$T(t) = T_0 + \frac{I^2 Rd}{kA}\left(1 - e^{-kAt/mcd}\right)$$

As time increases, the final temperature becomes

$$T_f = T(t \to \infty) = T_0 + \frac{I^2 Rd}{kA}$$

To see how this depends on the radius of the conductor, r, we may rewrite this as:

$$\Delta T = T_f - T_0 = \frac{I^2 Rd}{kA} = \frac{I^2 \rho \frac{L}{\pi r^2} d}{k 2\pi rL} = \frac{I^2 \rho\, d}{k 2\pi^2 r^3}$$

From this we conclude that the temperature increase due to current in the wire is:

(13.2) $$\Delta T \propto \frac{I^2}{r^3}$$

This means that if two wires have the same ΔT and one wire has twice the current of the other wire, the diameter of the larger wire must be a factor of $2^{2/3}$ = 1.59 times that of the smaller wire and the cross-sectional be a factor of $2^{4/3}$ = 2.52 larger.

Table 13.2 is a useful summary of wire gauges commonly used in residential wiring. (AWG stands for American Wire Gauge.) There is a general rule of thumb that you should not exceed 4.0 A /mm^2 in copper wire or 2.3 A/mm^2 in aluminum wire. Maximum currents obtained using this rule are indicated in Column 4. Since 12 AWG wire is typically used for 20 A circuits in homes, the last column, based on Eq. (13.3) probably presents more realistic values of maximum currents.

Table 13.2. Data for common copper wire sizes

AWG	diameter (mm)	Typical use	I_{max} (A) (4 A/mm^2)	I_{max} (A) Eq. (C.1) normalized to 20A for 12 AWG
6	4.12	electric stoves	53.3	57.0
10	2.59	water heaters, electric dryers	21.1	28.4
12	2.05	kitchen, dining room, bathroom, utility areas (best for most household circuits)	13.2	20.0
14	1.63	low current house-hold circuits (best for three-way lights)	8.3	14.2
16	1.29	Low voltage wiring: doorbells and thermostats	5.2.	10.0
18	1.02		3.3	7.0

Things to remember:

• In home wiring: black and red are hot, white is neutral, and bare or green is ground.

• Copper wire conducts better, is more flexible, and corrodes less than aluminum wire. Aluminum wire is less expensive.

• Larger wire must be used in circuits that draw more current. 6 AWG copper wire is used for stoves, 12 AWG wire is used for most household circuits.

13.5. Switches and Outlets

The purpose of switches, of course, is to start and stop the flow of current through circuits. Switches are characterized as single-pole or double pole, depending on whether one or two wires are connected. They are also termed single-throw or double-throw, depending on whether the switch is just open-closed or if the switch can transfer input current to two different output circuits. Schematic diagrams of various types of switches are illustrated in Fig. 13.11.

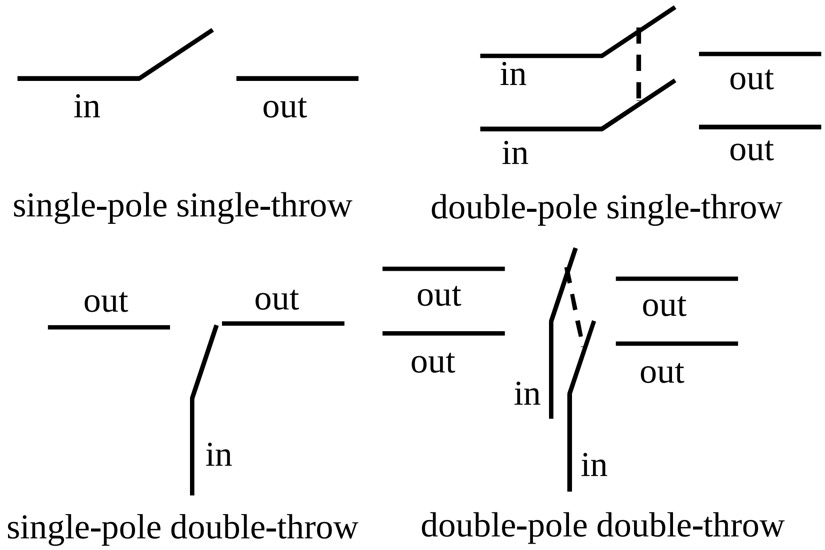


Figure 13.11. Common types of switches.

The most common switch used in house wiring is the single-pole single-throw (SPST) switch. It is often called just a single-pole or SP switch. SP switches are placed in the hot line to switch lights, outlets, or hard-wired appliances (such as bathroom fans) on and off. Below is a diagram of two ways switches can be wired. Note that the switch is enclosed in a receptacle for safety. Wire nuts are used to make a secure connection between wires. In these diagrams, hot will be black, neutral will be white, and ground will be gray.

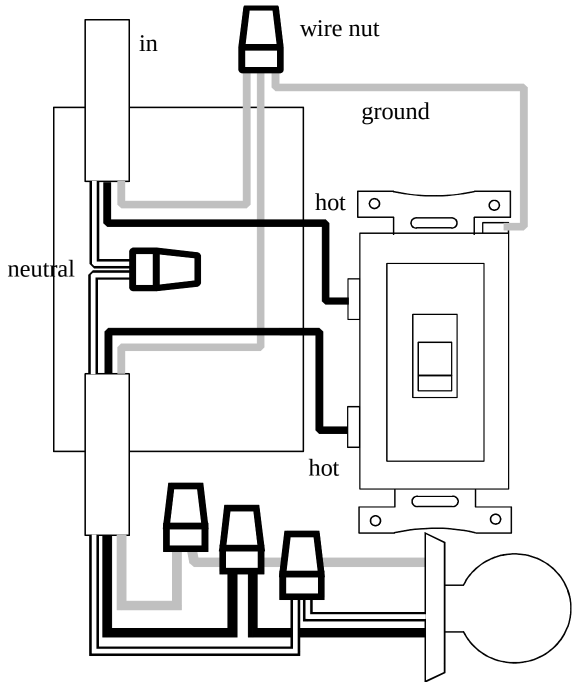


Figure 13.12. Wiring a simple light switch.

If two switches are in the same box, the inlet hot and neutral wires can be connected with wire nuts to both switches.

Although this is the simplest way to wire a switch, sometimes it is more convenient to connect the incoming power to the light. If this is the case, the switch can be wired as shown in Fig. 13.13.

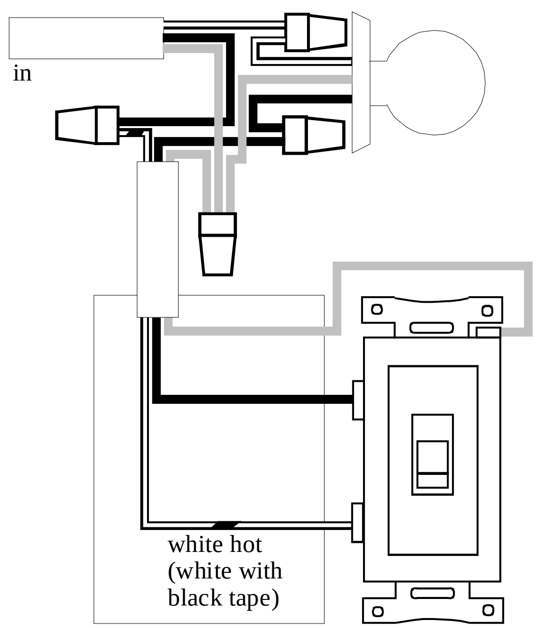


Figure 13.13. An alternative method of wiring a switch.

It is often desirable to have two different switches control a light, so that changing the position of either switch will change the on-off state of the light. In this case, simple SP switches will not suffice. The switches that are used are called “3-way switches.” These are just single-pole double-throw switches. A hot wire comes in to the common, or COM, connector of the switch. The switch then selects one of two output wires, called travelers. Two 3-way switches are connected to a light as shown in Fig. 13.14. Look through the diagram and be sure you understand how the circuit works.

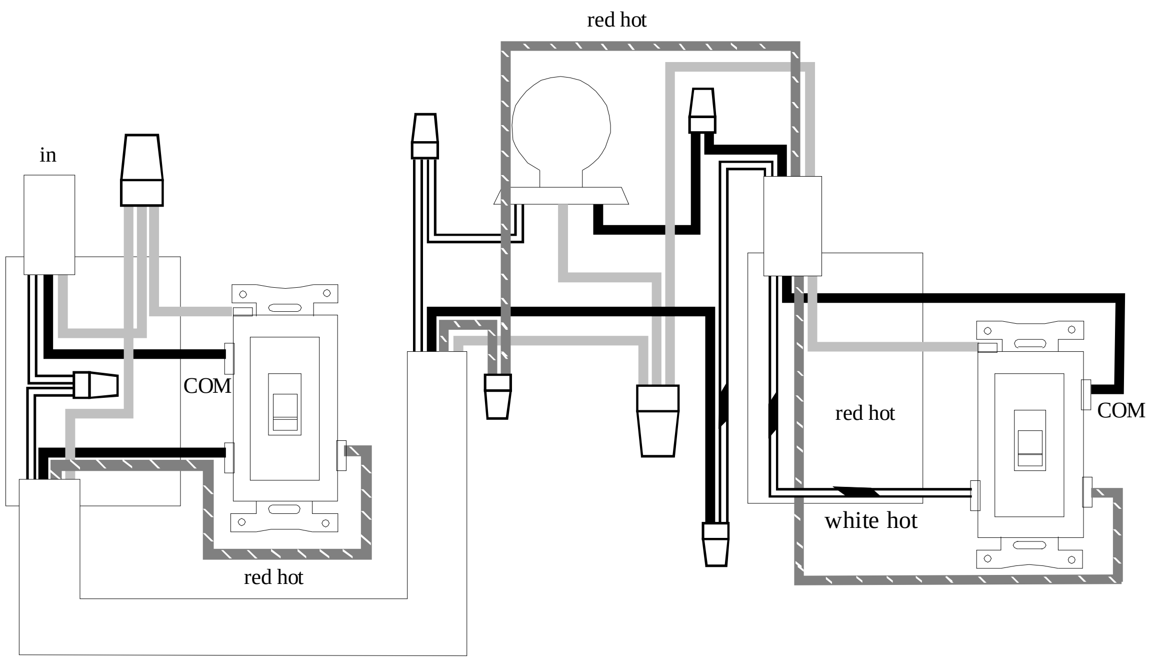


Figure 13.14. Wiring a 3-way switch.

Standard outlets have three different holes in them. The rounded hole is the ground, the long slot is neutral, and the short slot is hot. Just in case a plug is not in tight and something conductive is dropped on the plug, it is safest not to have the hot plug on top. Outlets should be installed in one of the configurations shown in Fig. 13.15.

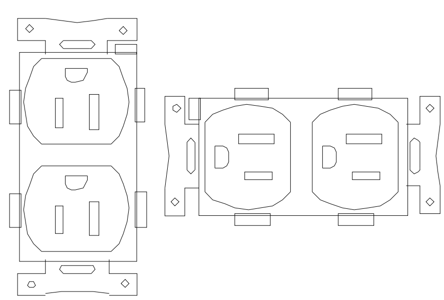


Figure 13.15. The proper directions to install outlets.

The connections on the side of the outlet nearest the hot slot are for the hot wires. These two connections are joined by a copper strip, so they are electrically equivalent to each other.

There are two common ways that a series of outlets can be joined in a circuit. These are called “series” and “parallel,” although the terminology is a bit misleading. Series is most common, as it is easier to install; however, parallel has the advantages that if a wire comes loose in an outlet box (which seldom happens), other outlets are not affected. These are depicted in Fig. 13.16 and Fig. 13.17.

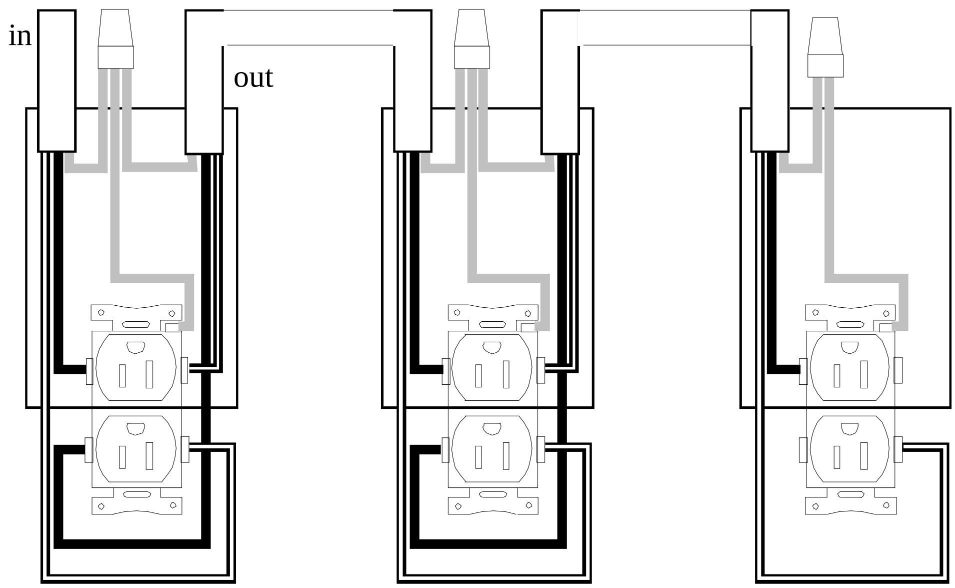


Figure 13.16. Outlets connected in series.

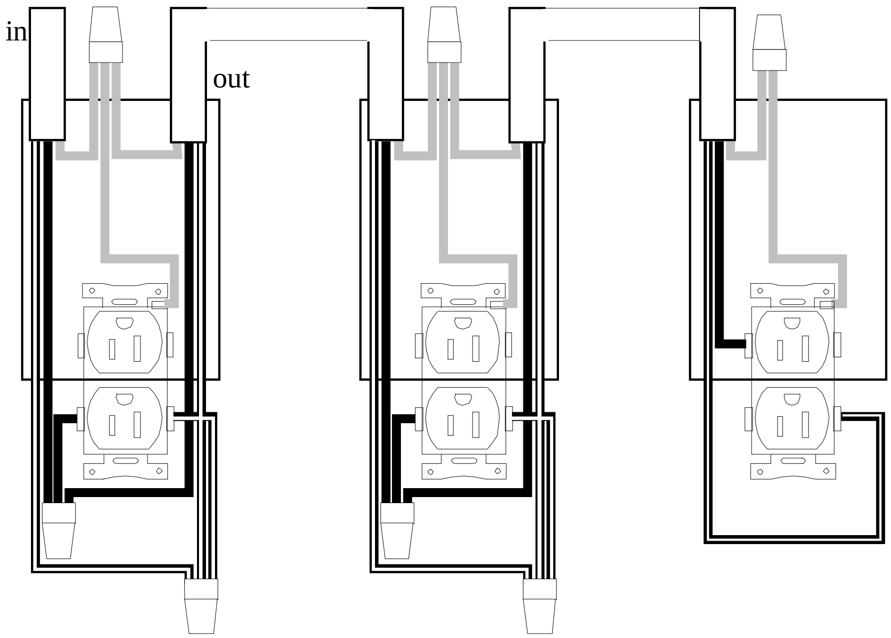


Figure 13.17. Outlets connected in parallel.

Things to remember:
• Switches are named by the number of wires being switched (poles) and the number of output options (throws).
• Switches are placed in the hot wire of a circuit.
• A white wire with black tape wrapped around it is hot (white hot).
• The hot connection of an outlet is the narrow one. Outlets should be installed with the hot connection down.
• Outlets can be connected in series and parallel.
• Be able to recognize a 3-way switch circuit, and series and parallel outlet circuits.

13.6. Safety Devices

To protect people from electrical hazards, a number of precautions are taking in wiring homes. The most important of these is simple grounding. If, for example, the insulation on the hot wire in an appliance breaks down and the case of the appliance is a conductor, the entire appliance becomes hot; that is, the case of the appliance is at 120 V potential. If the appliance is not grounded and someone touches it, current can flow to ground through the person. (Since water pipes and furnace ducts are usually good grounds, you must be especially careful with electricity around them.) This current can cause burns or disrupt the heart. If the appliance is grounded, however, the ground wire and the person both provide paths to ground. That is, the wire and the person are two resistors in parallel. Since the wire has a much smaller resistance than a human body, almost all the current flows through the ground wire.

If a hair dryer, for example, is operating normally, the same current flows into the hot wire of the circuit as flows out the neutral wire of the circuit. However, if current flows through a ground wire or a person to ground, the neutral wire has less current in it than the hot wire. A simple device called a GFCI or ground fault circuit interrupter compares the current in the hot and neutral legs of a circuit. If they ever become unequal, a switch, much like a circuit breaker, immediately opens. Do note, however, that if current flows from the hot wire through your body, and back to neutral, the GFCI will do nothing.

The heart of a GFCI is a differential transformer, a small transformer that produces currents in opposite directions from the hot wire and from the neutral wire. As long as the currents in the hot and neutral wires are the same, the net voltage in the third branch of the differential transformer is zero.

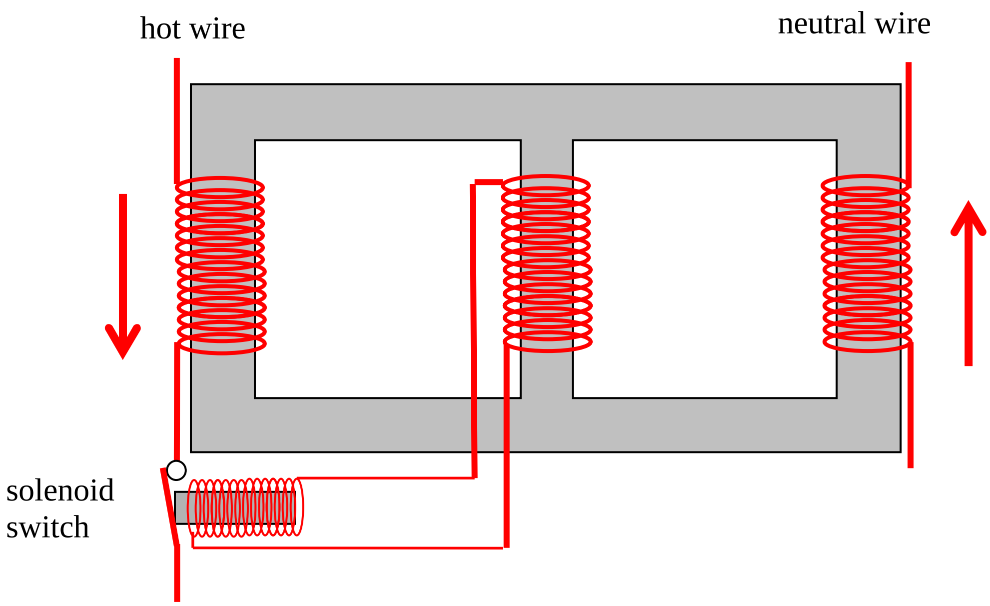


Figure 13.18. A Ground Fault Circuit Interrupter

GFCI's can be incorporated into circuit breakers; however, more often they replace a normal outlet in a circuit. A GFCI looks like a normal outlet, except it has a “TEST” and a “RESET” button on its face. GFCI's should routinely be installed in kitchens, bathrooms, and other locations where shock hazards are high. Since a GFCI may need to be manually reset if there are ground faults or power outages, it is important not to operate life-support equipment on GFCI protected circuits.

GFCI's are wired in series. Note that all outlets downstream from a GFCI are also protected (and need not be grounded), so you really need only one GFCI per circuit.

A similar device is an AFCI or arc fault circuit interrupter. If insulation wears in a circuit, an arc can result from a hot wire to a neutral or ground wire. Arcs don't draw enough current to shut off a normal circuit breaker, but they can produce dangerous levels of heat. AFCI's are typically incorporated into circuit breakers. Most building codes now require AFCI circuit breakers for any circuits that have bedroom outlets.

Things to remember:
• Know why grounding is important.
• Know what a GFCI is haw it operates. Know where GFCI's need to be installed and that they need to be installed in series.
• Know why AFCI's are used and where AFCI circuit breakers should be installed.

13.7. Waves: A Review

In the next sections we are going to learn more about electromagnetic radiation. To understand some of the terminology, however, we should first review a few things about waves.

The electrical field in an electromagnetic wave that varies in both space and time has the fairly general form:

(13.3) $$E(x,t) = E_0 \sin(kx \pm \omega t + \phi)$$

This is a sine wave moving in the $\mp x$ direction where:

E is the value of the electric field at x and t. It is measured in volts/meter (*V/m*).

E_0 is the amplitude (maximum electric field strength) of the wave in *V/m*.

k is called the "wavenumber." $k = \frac{2\pi}{\lambda}$ where λ is the wavelength. Its units are m^{-1}.

ω is the angular frequency of the wave. It is related to frequency, *f*, and period, *T*, by the expressions $\omega = 2\pi f = \frac{2\pi}{T}$. It is measured in *rad/s*.

ϕ is the phase angle. It tells us how the wave starts. That is, at *x=t=0* $E(0,0) = E_0 \sin\phi$. For our purposes, we can let $\phi = 0$.

Let's look at the wave in a little more detail to understand what these different quantities mean.

Equation (13.3) gives the electric field we measure at a position *x* and time *t*. This equation only tells us the magnitude of the electric field, not its direction (except for + and – signs). We know that sine varies between –1 and +1, so the largest value the wave can attain is the amplitude, E_0.

Let's draw a picture of a wave on a string at a particular time. (Waves on strings are easier to visualize than electric field strength. Just remember that electric field strength, *E*, works just like the displacement of a wave on a string.) We can think of this as a "snapshot" of the wave. For simplicity, let's choose the phase angle, ϕ, to be zero and take the snapshot at time *t=0*. What we plot then is just $y(x) = A\sin(kx)$ with A =5 mm and k = 2 rad/sec.

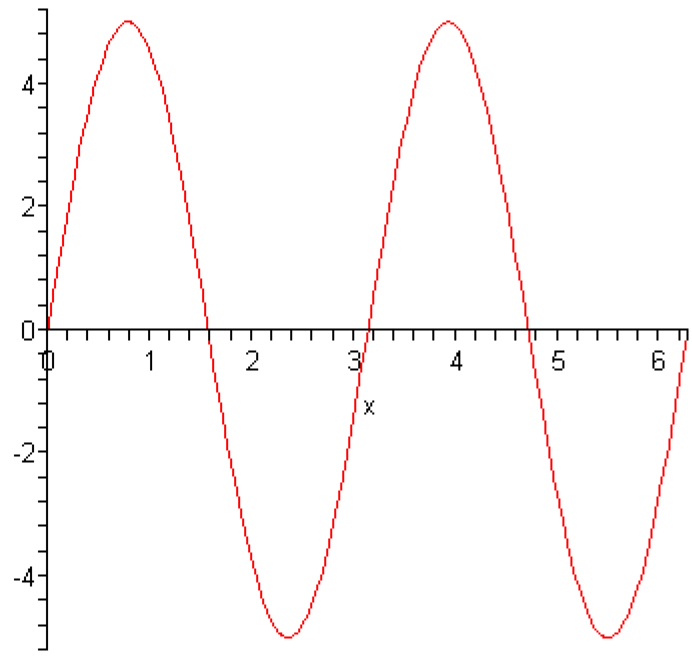


Figure 13.19. A snapshot of a wave on a string at time *t=0*.

The first thing we can do is ask where the peaks of the wave occur. In terms of angle we know those are at $kx = \pi/2$ and $\pi/2+2\pi$. In terms of *x*, the peaks are then at

$$x = \frac{\pi}{2k}, \quad \frac{\pi}{2k}+\frac{2\pi}{k}.$$

Since one wavelength, λ, is the distance between peaks, this means that :

$$\lambda = \frac{2\pi}{k} \quad or \quad k = \frac{2\pi}{\lambda}.$$

This gives us a very useful relationship between the wavelength and the wavenumber.

Now, let's take a snapshot of the wave just a little later. The function of the wave is now $y(x,t) = A\sin(kx-\omega t)$ with ω=20 rad/sec and *t* =0.01 sec. We know that the effect of adding this term is to translate the wave to the right by an angle $\omega t = 0.2$ rad. This is borne out in Fig. 13.20.

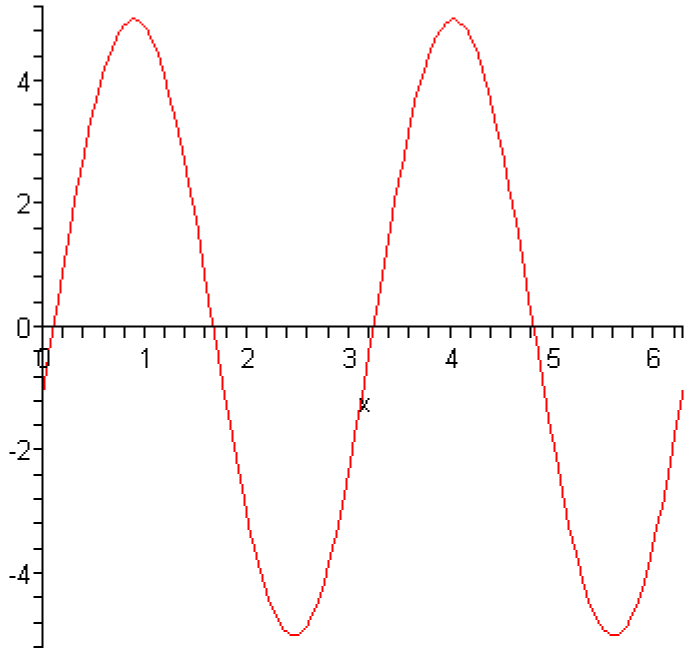


Figure 13.20. The sine wave of Fig. 13.19 a moment later.

As can be seen from these figures, the wave is moving to the right. We can tell how fast the wave moves by determining how far the wave moves in a time *t*. Since

$$kx - \omega t = k\left(x - \frac{\omega}{k}t\right)$$

the wave is translated to the right by a distance:

$$\frac{\omega}{k}t = vt \quad \Rightarrow \quad v = \frac{\omega}{k}.$$

From this, we can also see that a wave traveling to the lest must have the form $y(x,t) = A\sin(kx + \omega t)$.

Since waves are functions of both space and time, we can also choose a particular point along the wave, *x=0*, for example, and plot the wave at that point as a function of time. Since an oscilloscope shows a voltage as a function of time, we can think of this kind of graph as an "oscilloscope trace" of the wave. Mathematically, for a wave going to the right, the oscilloscope trace at x =0 is $y(0,t) = A\sin(-\omega t)$. This is illustrated in Fig. 13.21.

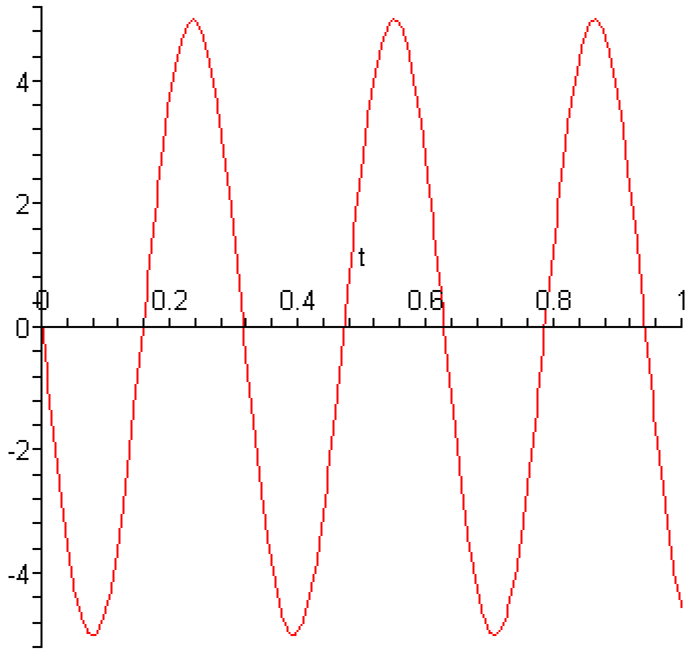


Fig. 13.21. An "oscilloscope trace" of the wave on a string.

Now we can again ask where the peaks of the wave occur. We know these are at angles $\omega t = \pi/2$ and $\pi/2 + 2\pi$. In terms of *t*, the peaks are then at

$$t = \frac{\pi}{2\omega}, \quad \frac{\pi}{2\omega} + \frac{2\pi}{\omega}.$$

Since one period, *T*, is the time between peaks, this means that :

$$T = \frac{2\pi}{\omega} \quad or \quad \omega = \frac{2\pi}{T} = 2\pi f.$$

Things to remember:

• $E(x,t) = E_0 \sin(kx \pm \omega t)$

• This wave moves in the $\mp x$ direction.

• E_0 is the amplitude.

• k is the "wavenumber." $k = \dfrac{2\pi}{\lambda}$.

• ω is the angular frequency $\omega = 2\pi f = \dfrac{2\pi}{T}$.

13.8. Maxwell's Equations and Radiation

In Lesson 11, we found that Maxwell's Equations could be in written in differential form as:

Gauss's Law of Electricity	$\nabla \cdot \vec{E} = \dfrac{\rho}{\varepsilon_0}$
Gauss's Law of Magnetism	$\nabla \cdot \vec{B} = 0$
Ampère's Law	$\nabla \times \vec{B} = \mu_0 \left(\vec{j} + \varepsilon_0 \dfrac{\partial \vec{E}}{\partial t} \right)$
Faraday's Law	$\nabla \times \vec{E} = -\mu_0 \dfrac{\partial \vec{B}}{\partial t}$

Let's consider a region of space where there are no charges and no currents, just fields. From Ampère's Law, we see that if there is a changing electric field, there must also be a magnetic field with curl. From Faraday's Law, we see that if there is a changing magnetic field, there must also be an electric field with curl. This suggests that changing electric fields with curl and changing magnetic fields with curl must go hand-in-hand.

The math is a little messy, but if you can take a step or two on faith, we can see what the mathematical implications of this are:

$$\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$$

$$\nabla \times \left(\nabla \times \vec{B} \right) = \mu_0 \varepsilon_0 \frac{\partial}{\partial t} \left(\nabla \times \vec{E} \right) = \mu_0 \varepsilon_0 \frac{\partial}{\partial t} \left(-\frac{\partial \vec{B}}{\partial t} \right)$$

$$\nabla \times \left(\nabla \times \vec{B} \right) = -\mu_0 \varepsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2}$$

The curl of a curl is a rather messy thing, but an identity from vector calculus (here you have to exercise the faith unless you want to work through the details – which you can do if you don't mind quite a bit of algebra) can simplify this relationship. This identity tells us

$$\nabla \times (\nabla \times \vec{B}) = \nabla (\nabla \cdot \vec{B}) - \nabla^2 \vec{B}$$

That doesn't look a lot better, but we do know from Gauss's Law of Magnetism that $\nabla \cdot \vec{B} = 0$. With that we can simplify the expression to

$$\nabla^2 \vec{B} = \mu_0 \varepsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2}.$$

While this may not look very promising, it's a beautiful thing to a mathematician – it's the wave equation. This is the equation satisfied by a wave with a velocity of

$$v = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}.$$

In about 1864, James Clerk Maxwell had just discovered the Maxwell term of Ampère's Law, $\nabla \times \vec{B} = \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$. Soon after he did, Maxwell went through this little exercise. He then plugged known values in for the constants and discovered that $v = c$, the speed of light. Maxwell believed he had discovered the secret of light. He was almost right; he discovered one secret of light: light is an electromagnetic wave.

Since most of us are not quite the mathematical genius that Maxwell was, let's go back and work through some of the details, making use of some facts we know about electromagnetic waves. From our study of the radiation of accelerating charges, we learned that far away from the source of radiation, the electric and magnetic fields are perpendicular and that $\vec{E} \times \vec{B}$ is in the direction the wave travels. We also found that the amplitude of the magnetic field is $1/c$ the amplitude of the electric field.

This suggest that a solution to Maxwell's equations would be:

$$\vec{E}(x,t) = E_0 \sin(kx - \omega t)\hat{y}$$
$$\vec{B}(x,t) = \frac{E_0}{c} \sin(kx - \omega t)\hat{z}$$

for a wave traveling in the $+x$ direction. Using the differential-operator forms of divergence and curl, we can see if these results do indeed satisfy Maxwell's equations.

$$div\,\vec{E} = \nabla \cdot \vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}$$

$$curl\,\vec{B} = \nabla \times \vec{B} = \hat{x}\left[\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}\right] + \hat{y}\left[\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}\right] + \hat{z}\left[\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right]$$
$$= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ B_x & B_y & B_z \end{vmatrix}$$

$$\nabla \cdot \vec{E} = \frac{\partial E_y}{\partial y} = 0 \qquad \rho = 0 \quad away \quad from \quad the \quad source$$

$$\nabla \cdot \vec{B} = \frac{\partial B_z}{\partial z} = 0$$

$$\nabla \times \vec{E} = \hat{x}\left[-\frac{\partial E_y}{\partial z}\right] + \hat{z}\left[\frac{\partial E_y}{\partial x}\right] = \hat{z}E_0 k \sin(kx - \omega t)$$

$$\frac{\partial \vec{B}}{\partial t} = -\hat{z}\frac{E_0}{c}\omega \sin(kx - \omega t)$$

$$\nabla \times \vec{E} = -\frac{kc}{\omega}\frac{\partial \vec{B}}{\partial t} = -\frac{\partial \vec{B}}{\partial t} \qquad as \quad c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

$$\nabla \times \vec{B} = \hat{x}\left[\frac{\partial B_z}{\partial y}\right] + \hat{y}\left[-\frac{\partial B_z}{\partial x}\right] = -\hat{y}\frac{E_0}{c} k \sin(kx - \omega t)$$

$$\frac{\partial \vec{E}}{\partial t} = -\hat{y}E_0 \omega \sin(kx - \omega t)$$

$$\nabla \times \vec{B} = \frac{k}{\omega c}\frac{\partial \vec{E}}{\partial t} = +\mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$$

Therefore, these functions do satisfy Maxwell's equations.

Things to remember:
- Be sure you understand the terms amplitude, wavenumber, angular frequency, frequency, and period.
- The electric and magnetic fields are perpendicular to each other and perpendicular to the direction the wave travels. The vector $\vec{E} \times \vec{B}$ points in the direction of the velocity.
- The amplitude of the magnetic field is $1/c$ times the amplitude of the electric field.
- The electric and magnetic fields are in phase. That is, if the electric field is large and a given point in space and time, the magnetic field is also large at that point.
- Know that $c = \dfrac{1}{\sqrt{\mu_0 \varepsilon_0}}$.

13.9. Electromagnetic Radiation and Radio Waves

Electromagnetic radiation can be produced with any arbitrary frequency. We can produce radiation at low frequencies by letting charges undergo oscillatory motion. Higher frequency radiation can be produced by oscillating electric circuits. Beyond what we can produce with circuits, we can produce even higher frequency radiation in atomic and nuclear transitions or by accelerating very high-energy particle beams. Whatever the frequency, we can deduce the wavelength by the important relation:

(13.4 wavelength-frequency relationship) $c = \lambda f$

where:

c is the speed of light.

λ is the wavelength.

f is the frequency.

This relationship is more or less intuitive. If twenty full wavelengths go past you in a second and each wave is 3 meters long, the velocity is 60 m/s. We can, however, derive the relationship from results of the previous section.

$$c = \frac{\omega}{k} = 2\pi f \frac{\lambda}{2\pi} = \lambda f .$$

Electromagnetic waves are often classified by their wavelength or frequency. The divisions between the types of radiation are rather hazy, but the following is a useful table, nonetheless.

Table 13.3. Electromagnetic Radiation

Name	Typical Source	Approximate Wavelength
Radio	Oscillating circuits	>10 cm
Microwave	Electronic devices	100 μm – 10 cm
Infrared	Atoms, molecules	700 nm – 100 μm
Visible Light	Atoms	400 – 700 nm
Ultraviolet	Atoms	1 – 400 nm
X-rays	Inner shells of atoms	1pm – 1nm
Gamma-rays	Nuclei	< 1pm

While each region of the electromagnetic spectrum is interesting for different reasons, we are going to spend some time considering radio waves as a special case. Radio waves can be created and detected by simple circuits that we can construct with the knowledge we have gained in this course. (Of course, many radio circuits are much more complicated that the ones we will study in this section.)

The basic principle of radio communications is to combine a signal with a carrier wave, broadcast the wave, receive the wave, and separate the carrier and signal once again. Therefore, we want to superimpose an audio wave or a digital signal on a radio wave for the wave to be useful for communication. The simplest way we can do this is to turn the wave on and off and send a digital signal. The earliest electromagnetic communications used Morse code in this fashion. More frequently, however, we want to change some characteristic of the wave at a signal frequency, ω_s. Usually the signal frequency is the frequency of a sound wave in speech or music. This frequency varies in time, but much more slowly that the radio wave varies in time. The process of changing the radio wave at the signal frequency is called "modulation." There are three principle types of modulation: amplitude modulation (AM), frequency modulation (FM), and phase modulation (PM).

Let us begin by taking a high-frequency sine wave as our basic carrier wave. That is, when there is a plain sine wave, no information is carried on the signal. Such a wave is shown in Fig. 13.22. A typical carrier wave has a frequency in the MHz range.

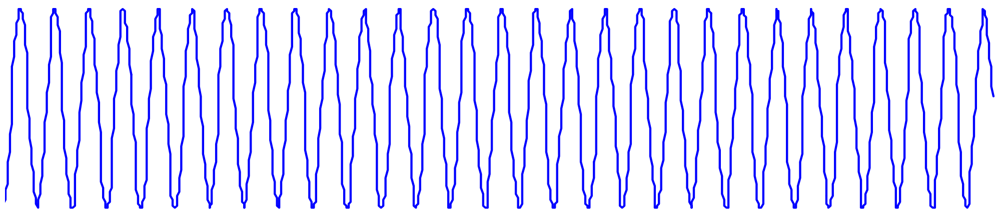


Figure 13.22. A high frequency carrier wave.

The information we wish to convey is contained in a wave of much lower frequency. Audio signals are in the 20 Hz – 20 kHz range.

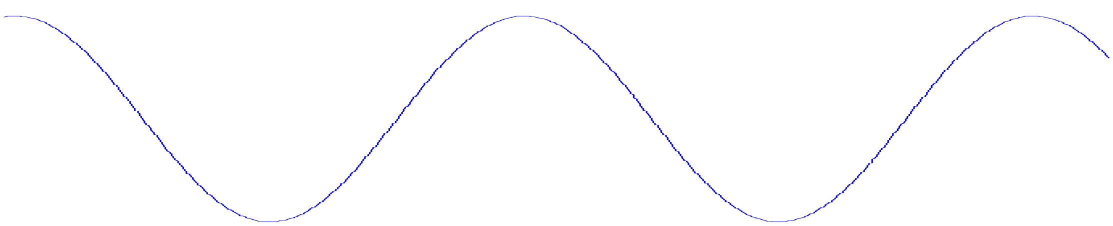


Figure 13.23. A lower frequency audio signal.

As the name implies, amplitude modulation changes the amplitude of the carrier wave as a function of time. Mathematically, we may write the modulated wave as:

$$E(x,t) = A\sin(\omega_s t)\sin(k_c x - \omega_c t) = A\sin(\omega_s t)\sin\left(\frac{\omega_c}{c}(x - ct)\right)$$

$$where \quad k_c = \omega_c / c.$$

$$E(0,t) = A\sin(\omega_s t)\sin(-\omega_c t)$$

A plot of this as a function of time for *x=0* is shown in Fig. 13.24 below.

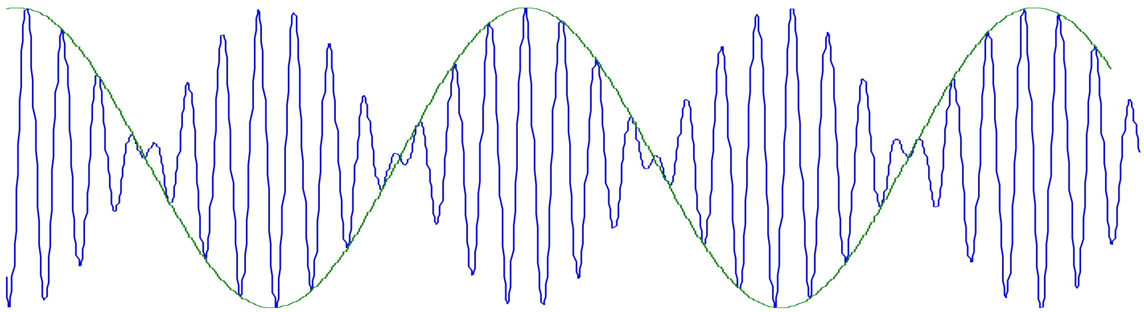


Figure 13.24. An AM wave with the audio signal superimposed for comparison.

The second method of modulating the wave is to change the phase as a function of time. In amateur radio, phase modulation is frequently used. PM waves are produced by a device called a **reactance modulator**. The mathematical and graphical representations of phase modulated wave are shown below.

$$E(x,t) = B\sin\left[\frac{\omega_c}{c}(x - ct) + A\sin(\omega_s t)\right]$$

$$E(0,t) = B\sin\left[-\omega_c t + A\sin(\omega_s t)\right]$$

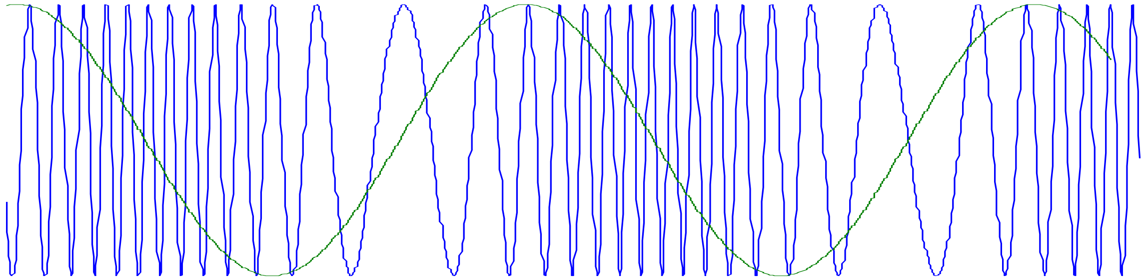


Figure 13.25. A PM wave with the audio signal superimposed for comparison.

The last kind of modulation, frequency modulation, is very similar to phase modulation because whenever we change the phase, we momentarily change the frequency as well. A frequency modulated wave has a higher frequency when the signal wave is large, and a lower frequency when the signal wave is small. It is illustrated below:

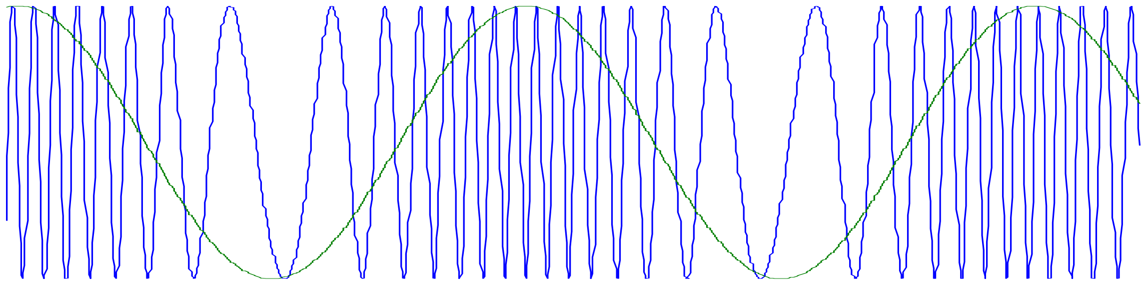


Figure 13.25. A PM wave with the audio signal superimposed for comparison.

This section is a little mathematical diversion. You can skip it, if you don't feel inclined to read it.

Mathematically, frequency modulation is a little more subtle. It seems that all we would need to do is write the frequency as the carrier frequency modulated by the source frequency:

$$E(x,t) \stackrel{?}{=} B\sin\left[\frac{\omega_c(1+A\sin(\omega_s t))}{c}(x-ct)\right]$$

But this does not work. The problem is that the wave at a specified time is given as the wave that would have been produced had the frequency remained constant over the entire interval [*0*,*t*]. By graphing the function above, you can demonstrate to yourself that the frequency actually gets higher and higher in the course of time. Instead, let's assume that we know the entire argument of the sine function – we'll call it φ – at some time. Since $\varphi = kx - \omega t$,

$$\varphi(t+\Delta t) = \varphi(t) - \omega_c(t)\Delta t$$
$$\frac{\varphi(t+\Delta t)-\varphi(t)}{\Delta t} \to \frac{d\varphi}{dt} = -\omega_c(t)$$

We haven't done anything to modulate the carrier frequency yet. But now we can simply consider the carrier frequency to be a base frequency, ω_0, modulated by the signal frequency. That is,

$$\omega_c(t) = \omega_0(1+A\sin\omega_s t)$$
$$\frac{d\varphi}{dt} = -\omega_0(1+A\sin\omega_s t)$$
$$\Rightarrow \varphi(t) = -\omega_0 t + \frac{A\omega_0}{\omega_s}\cos(\omega_s t)$$

To produce a radio wave, all we have to do is accelerate electrons back and forth along wires. The simplest form of transmission antenna is called a center-fed dipole antenna, as illustrated in Fig. 13.26. This type of antenna is typically constructed from two equal lengths of bare wire connected to a coaxial cable from an oscillating circuit. (A coaxial cable has a hollow cylindrical wire surround a normal cylindrical wire. Why doesn't the coaxial cable emit radiation?) The antenna radiates well if the length of each segment is chosen to be one-half of the wavelength of the radio wave, as measured in air.

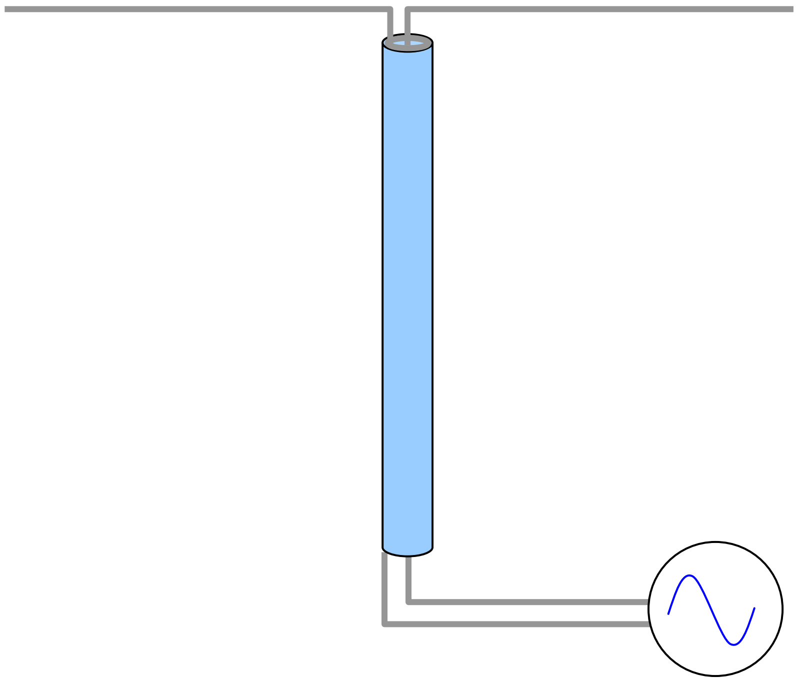


Fig. 13.26. A center-fed dipole antenna.

If the circuit is arranged properly, we can set up a standing wave on the antenna, in much the same way that we can set up a standing wave on a string. Note that, since current can not pass in or out of the ends of the antenna, the current at the ends must be zero, in analogy to the amplitude of the string in Fig. 13.27.

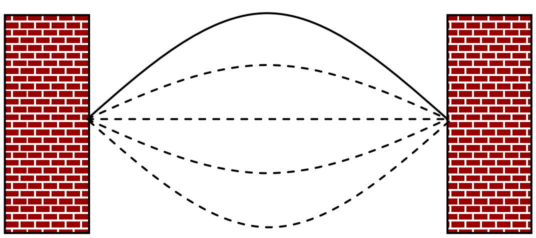


Figure 13.27. A standing wave on a string.

If we take $x = 0$ to be the center of the antenna and $x = \pm L/2$ to be the ends, then we may write for the current in the antenna

$$i(x,t) = i_0\left(\cos\frac{\pi x}{L}\right)(\cos\omega t)$$

where i_0 is the maximum current in the antenna and $\omega = 2\pi f$ is the angular frequency of the oscillating circuit. Furthermore we choose L to be one-half wavelength, so

$$L = \frac{\lambda}{2} = \frac{c}{2f}.$$

Here we have made use of the relation $c = \lambda f$.

To receive a radio wave, we can use a very similar antenna, called a half-wavelength dipole antenna. The electric field of the radio wave causes electrons to oscillate back and forth along the antenna. The antenna in the receiver circuit is just like an AC power supply. But there is one problem: a radio antenna is constantly being bombarded with thousands of signals from many different sources. We need to be able to tune a radio to a given frequency. To do this, we just use a series LRC circuit. The capacitor in the LRC circuit has a high impedance for low-frequency oscillations, and very little current will flow. The inductor has a high impedance for high-frequency oscillations. It's only when the carrier wave is at the resonance frequency of a circuit that much current can flow. The resonance frequency of the circuit is usually adjusted by changing the capacitance of a variable capacitor.

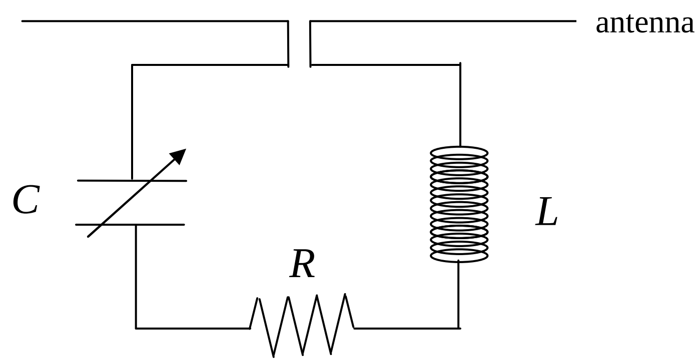


Figure 13.28. An LRC circuit for receiving radio signals.
The arrow through the capacitor indicates that it's a variable capacitor.

Of course, a resonating circuit by itself isn't sufficient to produce a sound on a speaker. The voltage, say across the resistor, must be amplified and the carrier wave separated from the signal in order to make a workable radio.

Things to remember:
• Know the electromagnetic spectrum. You don't need to remember the wavelengths listed in the table, but you should remember the names of the different types of electromagnetic radiation and their order in the spectrum.
• $c = \lambda f$.
• Radio communication requires us to modulate a carrier wave with a signal wave, transmit the wave, receive the wave, and then separate the signal from the carrier once again.
• Be qualitatively familiar with the three methods of modulating carrier waves: AM, FM, and PM.
• Center-fed dipole antennas are often used for transmission. Similarly, half-wavelength dipole antennas are often used for reception.
• The length of an antenna is generally on the order of one wavelength.
• A series LRC with a variable capacitor is used to tune the radio. The capacitor adjusts the resonant frequency of the circuit to match the transmission frequency.

13.10. Carrying Information on Electromagnetic Waves

With modern demands on sending and receiving more and more information in shorter and shorter times, it is very important to understand some of the limitations of electromagnetic transmission. Two important parameters in determining how much information can be put on waves are frequency and bandwidth.

1. Frequency: In order for a carrier wave to maintain its basic frequency so that it can be detected and tuned, the modulation frequency needs to be less than the carrier frequency. In other words, it is difficult to put more than about one bit of data on a full wavelength. Therefore, the higher the frequency – and hence the shorter the period and the shorter the wavelength – the more quickly data can be transmitted.

2. Bandwidth: Tuning circuits are not perfect. If a second signal with nearly the same signal as the carrier signal is picked up by an antenna, the tuning circuit will not be able to filter it completely out and the signal will be muddled. You have probably experienced two radio stations coming in to your radio at the same time. If a wide range of frequencies is available – either by regulation or for technical reasons – then multiple signals can be transmitted simultaneously. The difference between the maximum and minimum frequencies that can be used is called the "bandwidth." The number of signals that can transmitted at the same time without interference depends on the characteristics of the transmission and receiver circuits as well as upon the bandwidth. In general FM signals require more bandwidth than AM signals because the frequency of the FM signal is modulated.

Note that by extension of this meaning, the term bandwidth is also used for the number of bits per second that can be transmitted by any means.

Things to remember:
• The maximum rate at which information can be sent on a wave is about the same as the frequency.
• bandwidth is literally the frequency range over which signals can be transmitted. Bandwidth determines the number of signals that can be transmitted simultaneously.

13.11. Polarization

If we think of the radiation of a single point charge oscillating sinusoidally, we can easily deduce the direction of the electric and magnetic fields in the wave by using the methods of Lesson 10. The one thing that makes the analysis a little complicated is that we have to think of the motion of the source when the thread is emitted, not the motion of the source when the thread arrives at a field point. In Fig. 13.29 we wish to find the fields at point *P* die to a charge (we choose the nearest charge for convenience) oscillating in the antenna. Let's assume that the threads arriving at *P* were emitted from the source when the source acceleration was in the +*x* direction. The vector from the source to *P* is $\vec{R}$. We know that the direction of the electric field is then $\hat{R}\times(\hat{R}\times\hat{a})$, which is to the left. The direction of the magnetic field is $\hat{R}\times\hat{E}$, out of the . (Work these out yourself to be sure you agree.)

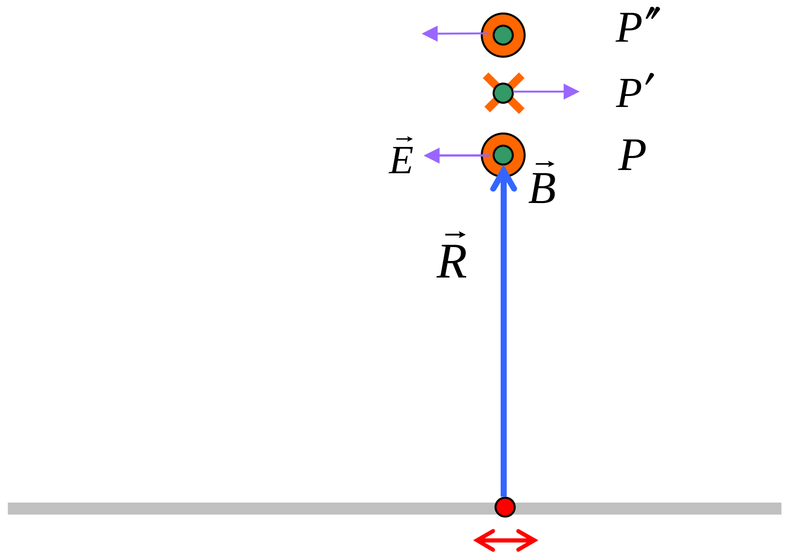


Figure 13.29. The fields of an oscillating source.

If we look at a point P' a little farther from the wire, so that the threads reaching that point were emitted when the oscillating source was accelerating to the left, both the electric and magnetic fields are reversed in direction. At a third point a little farther from the wire still, the field directions reverse once again.

From this figure, we see that the frequency of the electromagnetic wave is the same as the source's oscillation frequency and that the wavelength is $\lambda = cT = c/f$ as we already have observed. But most importantly, we note that the electric field is always oriented sideways. It gets larger or smaller, depending on position and time, but its direction is always sideways. Similarly the magnetic field is always directed in or out of the screen.

We define polarization to be the direction of the electric field in an electromagnetic wave. Unlike a normal vector direction, however, the polarization has two directions, such as up and down, sideways, in and out. This, of course is due to the fact that the direction of the electric field is oscillating in time.

Now what happens when we have many oscillators moving in different directions. This is what we have in light bulbs or light from the sun. In Fig. 13.30, we see light coming toward us from the sun. We know that the atoms in the sun that emitted the light oscillate randomly. However, we also know that the electric field is perpendicular to the direction of motion, so that there can be no component of the electric field in or out of the screen. Another way to say the same thing is that all the threads emitted by charges in the sun must lie in the plane of the screen. This is illustrated by the arrows drawn in all directions. We say that light from the sun is "unpolarized."

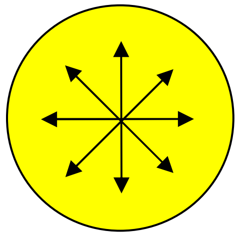


Figure 13.30. Unpolarized light from the sun.

Now let's think of sunlight reflecting off a lake. Light comes from the sun along the ray $\vec{r}_i$. When the light strikes the lake, the electric field of the light causes electrons in the water to oscillate, just as radio waves cause electrons in a receiving antenna to oscillate. The electrons in the water effectively become little transmitting antennas that re-radiate light to your eyes. These "antennas" emit radiation preferentially in the plane perpendicular to the line along which the water molecules oscillate.

Let's first consider light that is polarized in the horizontal direction, as shown in Fig. 13.31. The electric field of the sunlight causes electrons on the lake's surface to oscillate in and out of the screen, as indicated by the red circle on the water's surface. These electrons then radiate primarily in the plane perpendicular to their oscillation; that is, in the plane of the screen. Light from the water surface then strikes our eyes with polarization in the horizontal direction.

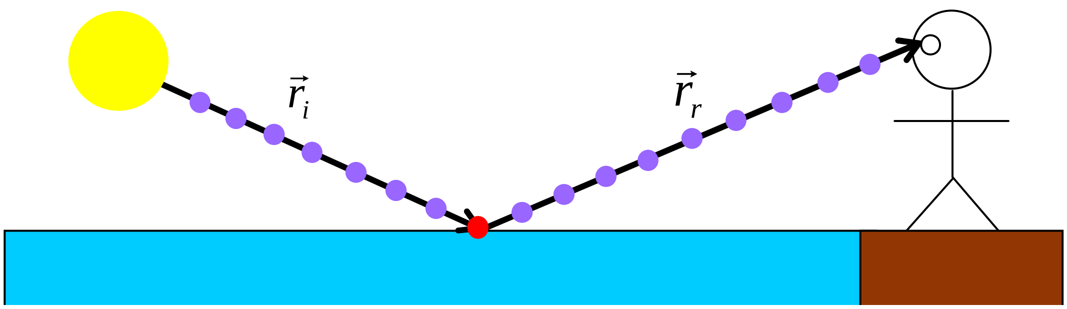


Figure 13.31. Reflected light polarized horizontally.

If light is polarized in the opposite direction – we'll call it non-horizontally – then the electrons in the water oscillate in the direction of the red arrow in Fig. 13.32. These electrons primarily emit light in the plane perpendicular to this arrow. This means that very little of the light polarized in this direction will get to our eyes.

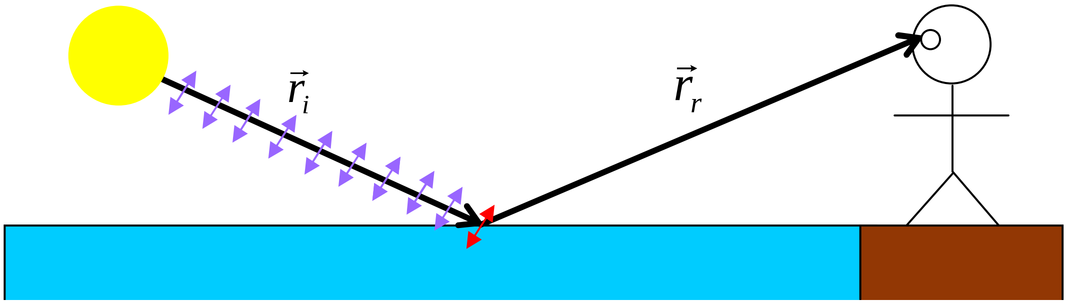


Figure 13.32. Very little light reflects when the polarization is in the opposite direction.

The bottom line is that the light that reaches our eyes is preferentially polarized in the horizontal direction. A more detailed analysis tells us that the amount of polarization is largest when the angle between in the incident and reflected rays is 90°.

Although it is believed that some animals are sensitive to polarization direction, we don't really notice the difference at all. If you have a pair of Polaroid sunglasses; however, you can detect polarization by rotating your lenses and see if the intensity of the light varies as you rotate.

Light can also be polarized by scattering, such as when light is scattered in the atmosphere. Light that comes from the sky at an angle of 90° from the sun is somewhat polarized.

A rather simple way to determine the direction of polarization in reflection and scattering processes is to consider the "polarization planes" of the incident and reflected (or scattered) rays. The polarization plane of a ray is the plane perpendicular to the ray; that is, the plane in which it is possible for the electric field of the light to point. The direction in which light is polarized is the intersection of the two polarization planes. This is depicted in Fig. 13.33.

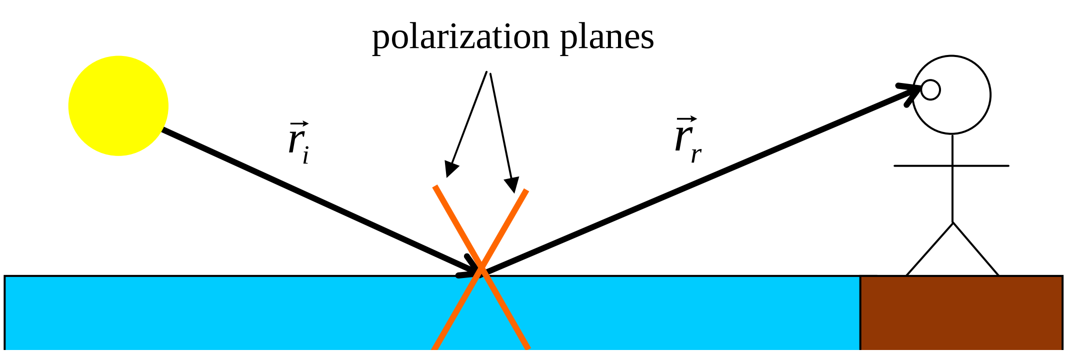


Figure 13.33. The intersection of the polarization planes on the surface tells us the preferred direction of polarization.

Although light can be polarized by reflection or scattering, a more effective means of polarizing light is by passing it through certain special materials. These materials come in two distinct types. The first type is a "birefringent" crystal, such as calcite. Because of peculiarities in the interaction of light with the crystal lattice, light that passes through calcite separates into two different rays with opposite polarization.

Replace with a photograph

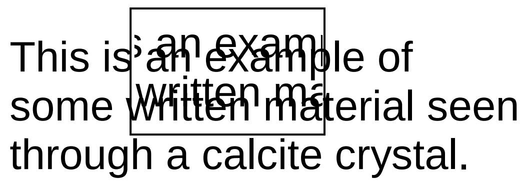


Figure 13.34. Writing viewed through a calcite crystal. The two sets of writing are polarized in different directions.

The second way material can polarize light is through selective absorption. Some materials with long molecules allow electrons to oscillate up and down along the molecules, but allow very little sideways motion of the electrons. The part of the wave that is polarized along the length of the molecules is absorbed, as the light energy is transferred to kinetic energy of the electrons. Light that is polarized in the opposite direction is not significantly affected. Such materials are called, "polarizers," "polarizing filters," or "Polaroid filters" after the company that developed them. To understand the physics of polarizers, we need a few facts:

1. Polarizers have an axis. The axis is defined to be the direction of polarization that can pass through the filter (as opposed to the direction that is absorbed).
2. The electric field vector of light entering a polarizer can be broken down into two components, the component along the axis and the component perpendicular to the axis.
3. The component of the electric field that is parallel to the axis is all that can pass through the filter.
4. The intensity of light is proportional to the energy in the electric field. As we learned in Lesson 6, the energy is in turn proportional to E^2. Hence, intensity is proportional to E^2.
5. When unpolarized light passes through a polarizer, its intensity is cut in half.

What happens if we have unpolarized light pass through two polarizers with their axes 90° apart? This situation is illustrated in Fig. 13.35.

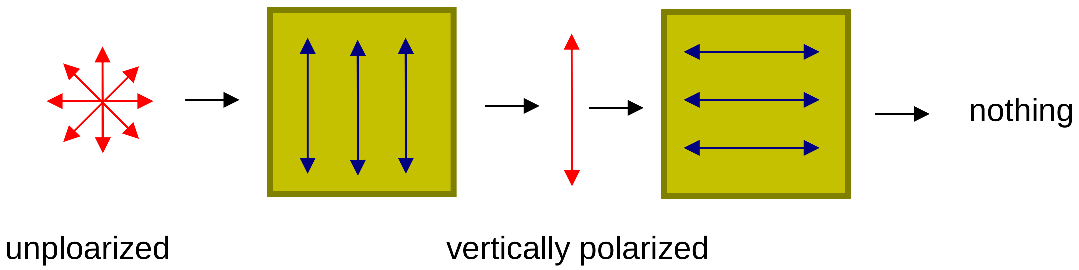
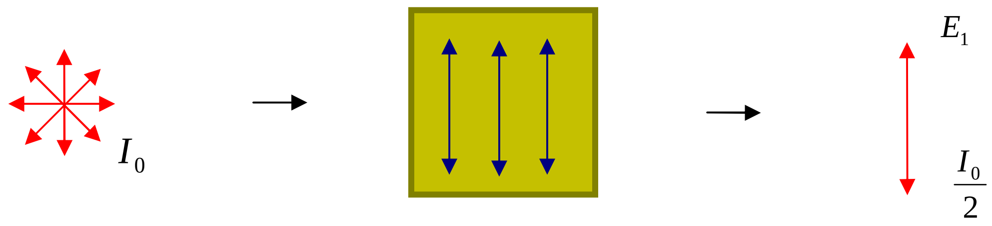


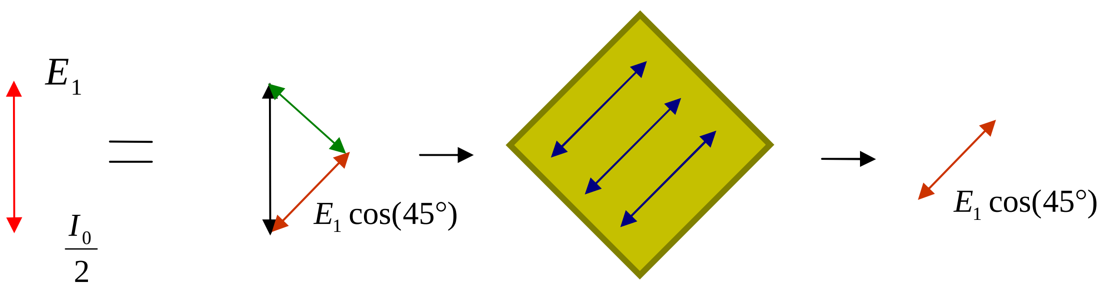
Figure 13.35. Light passing through crossed polarizers.

Example 13.1. Light passing through three polarizers.

Unpolarized light is incident on three polarizers. The axis of the second polarizer is at 45° with respect to the first and the axis of the third is at 45° with respect to the second. Note that if the middle polarizer is removed, the polarizers are at 90° and no light can pass through them at all.



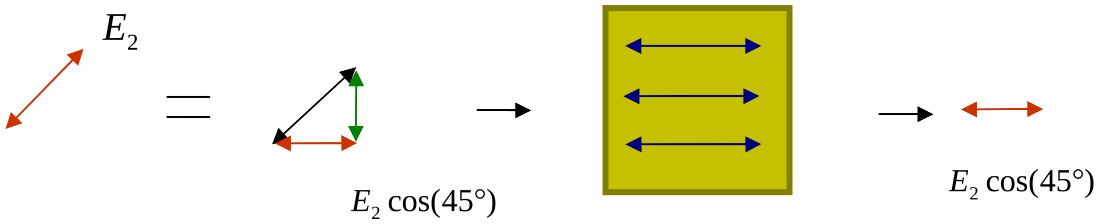
Half the initial intensity of the light is lost in passing through the first polarizer. After passing through this polarizer, the light is vertically polarized. We will call the length of its electric field vector (which we can represent by the length of the double-headed arrow) E_1.



After this stage, the ratio of the intensity to the original intensity, is:

$$\frac{I_2}{I_1} = \frac{E_2^2}{E_1^2} = \frac{E_1^2 \cos^2(45°)}{E_1^2} = \frac{1}{2}$$

$$\frac{I_2}{I_0} = \frac{1}{4}$$



So, finally we have:

$$\frac{I_3}{I_2} = \frac{E_3^2}{E_2^2} = \frac{E_3^2 \cos^2(45°)}{E_2^2} = \frac{1}{2}$$

$$\frac{I_3}{I_0} = \frac{1}{8}$$

After passing through all three polarizers, the intensity is 1/8 of the original intensity.

From this example, it is easy to see how the intensity of *polarized* light diminishes as it passes through a polarizer whose axis is at an angle θ with respect to the polarization direction. This result is known as Malus's Law:

(13.5 Malus's Law) $$\frac{I_{out}}{I_{in}} = \cos^2 \theta$$

Things to remember:

• Light can be polarized by reflection, scattering, birefringent crystals, or poalrizers.

• The direction of polarization for reflection or scattering is along the direction of the intersection of the polarizing planes of the incident and reflected (or scattered) rays.

• Polaroid filters cut out the component of polarized light that is perpendicular to the filter axis.

• Polaroid filters cut half the intensity of unpolarized light.

• The intensity of light is proportional to the square of its electric field.

• Malus's Law: $\frac{I_{out}}{I_{in}} = \cos^2 \theta$