Electricity and magnetism are really just two different faces of the same force. As in the electrostatic case, there are two kinds of magnetic “charges”, called poles. We call them north and south instead of positive and negative, but we still observe attraction between unlike objects and repulsion between like objects. We call a magnetic pole north if it points to the Earth’s magnetic north pole (this means, of course, that the Earth’s magnetic north pole must really be a south pole itself if it attracts other north poles). Just as we find the idea of an electric field useful, we also will introduce the idea of a magnetic field.

Magnetic field lines point from a north pole to a south pole, and are more closely spaced where the magnetic field is stronger. At any point in space, we will say that the direction of the magnetic field is the direction the north pole of a compass needle (itself a small magnet) would point if placed there.

There is one very important distinction between the electric and magnetic cases: no one has ever observed an isolated magnetic pole – magnetic poles are always found in pairs. Breaking a bar magnet into two pieces won’t give you a north pole and a south pole – it will instead give you two small pieces with north and south poles. These pairs of poles are known as dipoles, just as in the electric case. The difference is that electric monopoles (single charges like electrons or protons) are easy to find, and magnetic monopoles don’t appear to exist.

Magnetic Force on a Moving Charge

We would expect a magnet to feel a force in the field of another magnet, just as a charge feels a force in an electric field and a mass feels a force in a gravitational field. One of the things that illustrates the link between electricity and magnetism is the observational fact that a magnetic field can exert a force on an electric charge. For the charge to feel a force, it has to have a component of velocity perpendicular to the magnetic field. If the charge is stationary or is only moving along the magnetic field lines, it won’t be affected by the field. The strength of the force depends linearly on the strength of the field, the size of the charge, and the component of velocity perpendicular to the magnetic field. The direction of the force is perpendicular to both the velocity and the magnetic field. If you remember, this is the same kind of situation we had with torque. When we want to get a vector quantity (which torque was, and force is) from two other vectors (force and lever arm for torque, velocity and magnetic field for magnetic force), we will use the right hand rule to find the direction the vector points.

To find this direction, first point the fingers on your right hand in the direction of the particle’s velocity. Then, curl them in the direction of the magnetic field. Your thumb will be pointing in the direction of the magnetic force. This rule assumes that the charge itself is positive – an electron would be moving in the direction opposite to your thumb in the same situation.
The magnitude of the force is found by the formula below:

\[ F = qvB \sin \theta \]

where \( \theta \) is the angle between \( v \) and \( B \), \( B \) is the magnetic field (in units known as Teslas), \( q \) is the charge’s magnitude, and \( v \) is the velocity. As it turns out, the Tesla is a very large unit of magnetic field strength. The Earth’s magnetic field is about \( 10^{-4} \) T at the Earth’s surface which forms the basis for a more conveniently-size but non-SI unit of magnetic field strength called a gauss.

Because the force is perpendicular to the particle’s velocity, it has no component in the direction of the particle’s movement. This means that no work is done by this force (since Work = Force times displacement times the cosine of the angle between the two, and the cosine of 90° = zero). For this reason, the magnetic force won’t cause a change in the particle’s speed, but it can change its direction. You’ll probably see in the book that magnetic forces do no work. Actually, uniform magnetic fields do no work. A nonuniform magnetic field (where the field lines are not all parallel) can do work, but we won’t be looking at those.

What path will a charged particle take through a magnetic field? The force will always be perpendicular to the velocity of the particle, so it will make it curve. If the field extends over a large enough region (which depends on its magnitude and the particle’s charge and velocity), it will eventually make a circle. What will the circle’s radius be? The centripetal force on the charge is provided by the expression for the magnetic force above, so we get:

\[ qvB = \frac{m v^2}{r} \quad \text{so} \quad r = \frac{m v}{qB} \]

The radius of the circle made by the particle increases for higher mass or velocity and decreases for larger charges or magnetic fields. This effect is useful in determining the exact identity of a particle as it passes through a bubble chamber. These devices take advantage of the fact that a liquid which is free from impurities can be heated above its boiling point but is then super-sensitive to the introduction of new materials. If you’ve ever heard about microwave ovens and “exploding water”, this is the same phenomenon (which will require distilled water and a perfectly smooth coffee cup).
Superheated liquid hydrogen in the bubble chamber will show paths made by charged particles. If there’s a magnetic field going through the chamber, it will cause the particles to curve and the radius of the circles made by the particles can be measured.

We can use the same idea to identify molecules in a device called a mass spectrometer. It uses an electric field to give the charged particles a velocity (which is mass dependent – the particle’s charge multiplied by the potential difference will give its kinetic energy). A magnetic field will then bend the particles (also depending on mass and charge) and therefore provide a separation in space (we don’t have to let the particles make complete circles – we can just let them change direction so that they end up in a different location). Using this, we can distinguish one isotope of an element from another, since the charge is determined by the type of element but the mass depends upon the isotope since different isotopes have different numbers of chargeless (but not massless) neutrons.

**Magnetic Force on a Current**

If charges moving in a magnetic field feel a force, so will a current (since it’s just a bunch of moving charges). We know that charges have to move along the wire, rather than perpendicular to it. Therefore, we can take our expression for magnetic force on a charge and modify it slightly for currents. Where we have \( q \), we can replace that by \( q/t \) where \( t \) is the time it takes for a charge \( q \) to move through the wire. We’ll have to multiply something else in the formula by \( t \), though, because force divided by time isn’t force anymore. We get:

\[
F = \frac{\Delta q}{\Delta t} v\Delta t B \sin \phi = I \left( v\Delta t \right) B \sin \phi
\]

but velocity multiplied by time is just distance, in this case the length of the wire. Finally, we have:

\[
F = I L B \sin \phi
\]

We don’t even have to learn a new right-hand rule since the current \( I \) moves in the same direction as our charge \( q \) did previously.
Of course, since a current will be moving in a circuit, different parts of the circuit will feel different forces. In the first circuit below, the magnetic field is coming out of the page. Notice the directions of the forces, and use the right hand rule to see if you understand how to get them.

In the picture above, the magnetic field is coming out of the page (the circles with dots are supposed to represent the “point” of the arrow of the vector) and the current is moving clockwise. Notice that the force on every current segment is towards the inside of the current loop.

In the example below, the magnetic field is pointing from the left to the right. The currents which are parallel (or anti-parallel) to it feel no force, so \( F_1 = F_3 = F_5 = 0 \). The left and right sides of the loop feel forces in opposite directions. The force on \( I_2 \) points into the page (the circle with an “X” is supposed to be the feathered end of the arrow of the vector) and the force on \( I_4 \) points out of the page. If \( I_2 = I_4 \), the forces will be equal in magnitude and opposite in direction, giving us a pair of cooperating torques. The net action is to try to twist the loop so that side 4 comes out of the page and side 2 goes into the page.

This is one way to make an electric motor! A short way to remember how the loop will try to turn involves another right-hand rule. If you curl your fingers around in the direction of the
current, your thumb will be pointing perpendicular to the current loop (normal to it). The current-carrying loop is twisted in an effort to make its normal line up with the magnetic field.

What’s the magnitude of the torque? All we need to know is the distance between currents $I_2$ and $I_4$ so that we can find the lever arm of the magnetic force. Notice that even as the loop turns, the current remains perpendicular to the magnetic field, so the formula $F = I L B \sin \theta$ reduces to just $F = I L B$.

While the force remains the same as the loop turns, the torque does not. Once the loop has rotated 90° away from the configuration shown above (so that $I_3$ is coming out of the page and $I_1$ and $I_5$ are pointing into the page), the forces $F_2$ and $F_4$ will point towards the center of the loop, and will therefore exert no torque. We need to include a sinusoidal term (from the definition of torque) to take this into account. If $\phi$ is the angle between the normal of the current loop and the magnetic field, and $w$ is the width of the loop (distance between currents $I_2$ and $I_4$), our value for torque will be

$$\tau = I L B \left( \frac{w}{2} \sin \phi \right) + I L B \left( \frac{w}{2} \sin \phi \right) = I A B \sin \phi$$

where $A$ is the area of the current loop. If we imagine a coil of wire where the loop has been wound around a tube $N$ times, the torque will be $N$ times larger! Sometimes, we’ll find it useful to talk about torque applied by a magnetic field in terms of the strength of the field and the strength of a magnet twisting in it. A current loop acts like a magnet in some ways, though, so we can define a magnetic moment for the coil as just the product of the current, loop area, and number of loops:

$$\mu = N I A$$

The direction of $\mu$ is the same as the normal of the current loop(s).

We can build a motor by exploiting this effect. Even though there will be no torque produced when the loop is lined up with the magnetic field, the loop’s momentum will carry it through that dead spot. As shown in your book on page 635, the connection between the rotating current loop and the source of emf is arranged so that the process repeats every 180° to keep the loop turning.

### Magnetic Field of a Current

With the connections we’ve seen so far between electricity and magnetism, you may not be surprised to find that a current produces a magnetic field as well as responds to one. The magnetic field curls around the current-carrying wire in the same direction as the fingers of your right hand (surprised?) if you point your thumb in the direction of the current. To find the magnitude of the magnetic field, we realize that it should increase when the current increases but it should also decrease with distance from the wire. That’s about as close as we can get to deriving this particular formula, so here it is:
\[ B = \frac{\mu_0 I}{2\pi r} \]

where \( \mu_0 \) is known as the **magnetic permeability of free space**, \( I \) is the magnitude of the current, and \( r \) is the distance between the wire and the point you’re measuring \( B \). The value of \( \mu_0 \) is \( 4\pi \times 10^{-7} \) T m/A.

Because of this effect, we’ll see currents affecting the motion of moving charges (because a magnetic field affects the motion of a moving charge), other magnets, and things that look like magnets (other current-carrying wires and loops). Now that we know how to find the \( B \) produced by a current, we can just substitute it in where we need it. How will two current-carrying wires interact? We can look at the magnetic field produced by one and then find the force on the current carried by the other:

Notice that the magnetic field \( B_1 \) produced by current \( I_1 \) is pointing up at the location of current \( I_2 \). The force on current \( I_2 \) is perpendicular to both \( B_1 \) and \( I_2 \) and acts to push it directly **away** from the other current \( (I_1) \). If we reversed the direction of either \( I_1 \) or \( I_2 \) (but not both), the force would tend to push the two wires together rather than apart. When you think about it, this is weird – here we have like currents attracting and unlike currents repelling. This is the opposite of everything else we’ve seen so far, but that’s the way it is. Since we know the expression for the force a current feels in a magnetic field, and we know the expression for the magnetic field produced by a current, we can combine them to find out how hard these two push (or pull) on each other. We get (force per unit length of wire)

\[ \frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r} \]

In fact, this expression is used in the definition of the ampere (since it’s hard to count electrons as they go by!). An ampere is defined as the current in each wire when the two wires are one meter apart, one meter long, and pull on each other with a force of \( 2 \times 10^{-7} \) N.
Because of the distance-dependence of the force, a symmetric loop of current can be pulled towards or pushed away from a current-carrying wire if the parts perpendicular to the field are at different distances from it, as below:

In the drawing above, only the force between $I_5$ and $I_1$ and the force between $I_5$ and $I_3$ will be important in trying to understand which way the current loop will move. Only the magnetic field of $I_5$ is shown (circled dots and x’s). Since $I_1$ and $I_3$ are on the same side of $I_5$, they’ll both feel a magnetic field going into the page. Their currents are in opposite directions, though, so the forces they feel will also be oppositely directed ($I_1$ away from $I_5$, $I_3$ towards it). The magnitudes will be different, though, because the magnetic field of $I_5$ drops off with distance as we’ve seen before. The field will be weaker at $I_3$ than at $I_1$, therefore the force on $I_1$ will be larger than the force on $I_3$, and the loop will move away from the line of current.

The magnetic field produced by the current in the loop itself is shown below. By the right hand rule (one of them, anyway) you can put your thumb in the direction of the current and curl your fingers around the wire and see that the magnetic field points into the page inside the loop and out of the page everywhere else.
If we want to know the strength of the magnetic field at any point, we’ll have to add up the contributions from all of the wires. What’s a little easier (and generally more useful) is to imagine a circular loop just like the one above and ask what the magnetic field is at the center of it. For a coil with \(n\) turns per unit length carrying a current \(I\), the magnetic field at the center is

\[ B = \mu_0 I n \]

This coil, wrapped like a garden hose on a hanger, turns out to be very useful. We’ll call a collection of wire wound in loops like this a **solenoid**. Interesting things happen when the length of the solenoid is much longer than its diameter (think about a long, thin pipe wrapped with wire). The magnetic field inside starts to become nearly uniform (kind of like the electric field in a parallel-plate capacitor) and the field outside the solenoid is very small in comparison. At the ends of the solenoid, things get a little more complicated (just as we get edge effects in a parallel-plate capacitor), but if we stay near the center, we can say the field inside is uniform and large and the field outside is zero.

The interesting thing about this is that we’ve now made a magnet (called an **electromagnet**) that we can turn on and off with the flip of a switch. We can change the direction of the current going through it and the poles will swap ends. This effect is central to the operation of a **relay**, which uses an electromagnet to pull a switch on a spring open or closed. The switch can control a very large amount of current. This way we can use a small amount of current (in the electromagnet) to control a large amount of current. This is how the solenoid in your car can use the small amount of electricity passing through the thin wires in your steering column to control the flow of the very large amount of current needed to power your starter motor and get the engine running.

**Ampere’s Law**

Just as we had Gauss’ law which described the electric field all over an arbitrary surface which may or may not enclose a charge, we have a law that tells us about the magnetic field around a loop which may or may not go around a current. To clarify, Gauss’ law told us about the electric field on any **surface** we choose to imagine which is a boundary for some **volume** that may enclose some charge. **Ampere’s law** tells us about the electric field around a **line** which forms the boundary for some **surface** through which current may be flowing. Specifically, we get

\[ \sum B \parallel \Delta l = \mu_0 I_{enc} \]

Laws like this which give us fields in terms of boundaries and the regions they contain are very deep, and very powerful. Ampere’s law tells us that if we create some arbitrary closed loop, we can break the loop up into many tiny pieces, each of which is approximately a straight line. The component of the magnetic field parallel to that line segment, multiplied by the tiny line segment’s length, should be added up all around the closed loop. When we’re done, we’ll get just \(\mu_0\) multiplied by the current flowing through the surface of the loop.
In yet another similarity between the two laws, we’ll find ourselves looking at curves of high symmetry (circles, usually) just as we spent most of our time with Gauss’ law looking at surfaces of high symmetry (spheres). If we draw a circle of radius \( R \) around a long wire carrying a current of \( I \), we will notice that \( B \) is always parallel to our tiny little line segments of the circle (it’s because \( B \) wraps around the wire and always points tangentially, never having a radial component). Since all points on the circle are the same distance from the wire, the magnitude of \( B \) is always the same. Using Ampere’s law, we get

\[
B \left( 2\pi r \right) = \mu_0 I_{enc} \quad \text{so} \quad B = \frac{\mu_0 I_{enc}}{2\pi r}
\]

which should look familiar. This law is valid for a magnetic field which does not change as time goes on (steady-state current).

**Ferromagnetism**

You’re probably aware that magnetism created by electric currents moving through wires is not the only kind that exists. Permanent magnets have been around essentially forever, first noticed by humans as lodestone, a naturally occurring magnet. What you may not know is that these two kinds of magnetism are very closely related. The source of a permanent magnet’s magnetic field comes from the electrons making up the material. They move around the nuclei of the atoms they belong to in a way that mimics the flow of current (which is just charges moving in a circuit anyway). We can think of them as creating tiny current loops as they travel. Moving around the atom like that also tells us that they have orbital angular momentum, which we saw last semester. The Earth has orbital angular momentum around the Sun. The Earth also has spin angular momentum since it is rotating on its axis. Electrons, too, have spin angular momentum.
(but we don’t think of them as “spinning” on an axis – this is just one of the weird things we have to accept when we look at quantum mechanics later), and this also causes a magnetic effect. The overall effect is to make each atom like a tiny bar magnet (for some materials). If (as is common) all the magnets are pointing in random directions, the material as a whole won’t appear to be magnetic.

In some substances like iron, though, there is a tendency for very large groups of atoms to all line up in the same direction, forming **domains**. These little neighborhoods of cooperating tiny magnets act like much larger magnets. If a material like this is placed in a magnetic field, the domains that are already pointing that way will tend to grow at the expense of other domains that aren’t aligned with the field. A strong magnetic field can then magnetize a piece of iron (or a very small region in the thin coating of magnetic powder on the surface of a cassette, videotape, or computer disk).

If the magnet is heated, the random thermal motions of the atoms can overwhelm the cooperative effects of magnetism and let the atoms reorient themselves in essentially random directions. This is why you don’t heat a magnet with a torch unless you want to remove its magnetism. Similarly, jarring the domains by dropping the magnet or hitting it with a hammer will do the same thing.

**Induction**

There is yet another effect linking magnetic fields and currents, and that is **magnetic induction**. Imagine connecting a wire to an ammeter. Certainly, no current will flow since we don’t have a source of emf. If the connecting wire is in the shape of a coil, we still don’t see anything happen since there is no source of emf. If we pass a magnet through that coil, however, we will see a current appear! As soon as the magnet stops moving, though, the current disappears. If we remove it from its new position inside the coil, we’ll see a current (in the other direction) again! As soon as the magnet is safely out of range, the current disappears.

We see that it’s not the **presence** of the magnet, it’s the **change** in the magnetic field through the loop. That’s why moving the magnet in causes a momentary current which disappears when the magnet stops inside the loop. Remove it, and you’ve made a change again, so there’s another brief current.

We can do one more trick with the magnet – if we park it in the coil, the current will quickly disappear. If we change the **size** of the coil, we’ll see a current flow even while the magnet is stopped. The important thing turns out to be the change in **magnetic flux**. We looked at electric flux when we studied Gauss’ law. It was just the component of the electric field perpendicular to a surface multiplied by the area of that surface, and it gave us the charge enclosed (divided by $\varepsilon_0$). The magnetic flux is defined the same way, except we’re using the magnetic field, of course. The definition of magnetic flux is therefore

$$\Phi = B A \cos \theta$$
where $\Phi$ is the flux, $B$ is the magnetic field, $A$ is the area, and $\theta$ is the angle between the magnetic field and the normal to $A$. When this quantity changes, an emf is generated, and the magnitude of the emf is given by Faraday’s law of induction:

$$\varepsilon = -\frac{\Delta \Phi}{\Delta t}$$

Notice the negative sign; the emf is always directed so that the current it produces generates a magnetic field which opposes the change in the existing magnetic field. If the field is decreasing, an emf will be produced causing a current which adds to the magnetic field, trying to keep it constant. If the field is increasing, the current will give a magnetic field which opposes the original field. This method of finding the direction of the emf is known as Lenz’s law, and is a direct result of the conservation of energy.

One of the standard examples of this phenomenon is a resistor (or light bulb) connected to two conducting rails which lie in a magnetic field. Until the rails are connected by another conductor, the circuit is open. What will happen if they’re connected (no battery is in this circuit, remember) and the bar connecting them is moved?

In the drawing below, we have a magnetic field pointing into the page and a moving bar which completes the circuit. As the bar moves to the right, the magnetic flux increases (the magnetic field is assumed constant, but the area enclosed by the conducting loop is increasing as the bar moves to the right. If the induced emf is going to give us a current which causes a magnetic field which fights the change, the magnetic field needs to be pointing out of the page. If the field is pointing out of the page, the current must be circulating counterclockwise (right hand rule again).
What will the magnitude of the emf be? \( B \) isn’t changing, but the area is. The change in area per second is just the distance between the rails multiplied by the velocity of the connecting bar. The change in flux per time is therefore just

\[
\frac{\Delta \Phi}{\Delta t} = B L v
\]

where \( L \) is the distance between rails. Let’s say the rails are 1 m apart and the bar is moving with a velocity of 30 m/s. What will the induced emf be if the magnetic field = 0.2 T? Just multiply 0.2 T * 1 m * 30 m/s and we get 6 T m²/s or 6 Volts. If we know the resistance of the circuit, we can calculate other things. For example, if it’s 3 \( \Omega \), the current must be \( 6 \text{ V} / 3 \text{ \Omega} = 2 \text{ A} \). How much power is being dissipated in this circuit? Power = \( I V \), so it’s 2 A * 6 V = 12 Watts.

We can also find something else interesting: what force is necessary to keep the bar moving at a constant velocity? Remember that without a magnetic field, and ignoring friction, we wouldn’t need to apply a force to keep it moving because nothing would be acting to decelerate it. The magnetic field is causing something like friction and we have to apply a force to the bar to keep it going. How big is the force? Well, take advantage of the fact that we know that force and power are connected by \( P = F v \). We know \( P \) and \( v \) already, so just divide and find \( F \): 12 W / 30 m/s = 0.4 N. Not coincidentally, we can find the same answer using the force on a current we found in the last chapter:

\[
F = I L B \sin \theta
\]

\( B \) is perpendicular to the current in the bar, so \( \sin \theta = 1 \) and we get \( F = I L B = 2 \text{ A} * 1 \text{ m} * 0.2 \text{ T} = 0.4 \text{ N} \) again. We’ve done this two ways – first, using conservation of energy, and then using the magnetic force on a current. We’ll pick which one we want to use based on the specifics of a given problem.

**The Electric Generator**

Yet another example of the interplay between electricity and magnetism can be seen when we discover that the motor (which turns electricity into mechanical work) can work backwards as a generator (turning mechanical work into electricity). If we have a conducting loop in a magnetic
field and rotate it, we change the magnetic flux going through it by changing the angle between the magnetic field and the loop’s normal vector. This will cause a current to flow in the loop. We can use our earlier expression for the emf in the loop (or a coil of \( N \) loops) here:

\[
\varepsilon = -\frac{\Delta \Phi}{\Delta t} = -\frac{\Delta (NBACost)}{\Delta t}
\]

In this formula, \( N, B, \) and \( A \) should all be constant in time as long as the magnetic field is constant and the coils aren’t unwinding themselves or something. We can then take these three terms outside of the “delta” and see that only the change in the cosine of the angle with time is left. If the coil is turning at a constant angular velocity \( \omega \), that is related to the angle between \( A \) and \( B \) at any given time \( t \) by \( \theta = \omega t \) (assuming that the angle \( \theta \) was zero when \( t \) was zero). Now we have an expression for the emf that looks like

\[
\varepsilon = -NBA\frac{\Delta Cos(\omega t)}{\Delta t}
\]

Now all we have to do is figure out what the last part of this means. How does this cosine term change in time? There are two ways we can handle this: one, we can just realize that this is one of the things that calculus was made for and consult a book to find that

\[
\frac{\Delta Cos(\omega t)}{\Delta t} = -\omega Sin(\omega t)
\]

The other way we could handle this is to plot this cosine term vs. time and look at the slope at every point along the curve. If we make a careful study of the value of \( t \) and the slope of the cosine term at that particular value of \( t \) for a whole period, we’d see that we were really plotting \(-\omega Sin(\omega t)\) vs. \( t \). Either way, our final answer for the emf generated is

\[
\varepsilon = NBA\omega Sin(\omega t)
\]

where \( \omega \) is measured in radians per second. Notice that the emf is proportional to the strength of the magnetic field (stronger magnets are better), the size of the coil (larger is better), the number of loops of wire in the coil (more = better) and the angular speed of rotation (again, faster is better). We want a generator with lots of large loops of wire spinning quickly in a large magnetic field. Also, though, we have a term that oscillates in time. The emf is going from a maximum (where \( \omega t = \pi / 2 \)) to zero (\( \omega t = \pi \)) to a minimum (\( \omega t = 3\pi / 2 \)) through zero again (\( \omega t = 2\pi \)) and back to a maximum, completing the cycle. For convenience, we’ll sometimes write the time-varying emf as the product of the sine term and the maximum (peak) emf:

\[
\varepsilon = NBA\omega Sin(\omega t) \text{ or } \varepsilon_0 Sin(\omega t)
\]
The voltage is changing polarity as the coil rotates, making the wires connecting the generator to whatever we want to power flip back and forth from positive to negative through zero. If we have a purely resistive circuit attached (we’ll see later what complications arise when this isn’t true), the current will change as the emf changes, switching from a positive maximum through zero to a negative minimum (i.e., changing direction), etc. This is called, unsurprisingly, an alternating current. In the US, the power grid is based on generators operating at a frequency of 60 Hz (remember that frequency $f$ is connected to angular frequency $\omega$ by $\omega = 2\pi f$). If the other variables are fixed, we get a dependable emf between the two sides of an electrical outlet. The voltage across an “ordinary” wall outlet is about 120 V$_{\text{rms}}$ in the US. What’s the peak voltage? It must be $120 \text{ V} \times \sqrt{2} = \text{about 170 V}$. 

Drawing power from a generator isn’t free. We saw that a bar sliding along conductive rails in a magnetic field generates an emf (and therefore a current) but also experiences a force which tends to stop its movement. The electrical power produced by the bar has to be supplied by the mechanical force keeping the bar moving. The generator is no different – the resistors in the circuit which the generator is driving (known as the load placed on the generator) will dissipate energy as heat, and that energy needs to be supplied by whatever is rotating the generator’s coil. For a perfect generator delivering $P$ watts of power, the load will exert a torque on the coil which is found from $P = \tau \omega$ and is directed in such a way as to try to stop the coil from rotating (if it went the other way, the slightest resistance would cause the generator to increase its rotational speed, violating the conservation of energy). This torque is called the countertorque and is produced by the action of the magnetic field on the current circulating in the loop.

The similarity between a generator and an electric motor means that we’re going to see a similar effect in a motor. As the motor rotates under the influence of an applied emf $V$ (like the 120 V$_{\text{rms}}$ coming from the wall), the coil in the motor will act like a generator and produce its own emf $\varepsilon$ which will oppose the applied emf. This is known as the back emf or counter emf, and it makes the motor act like a smaller generator wired up to oppose the primary generator. This back emf depends on the rotational speed of the motor and the load on it. The current $I$ required by the motor will be

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

where $R$ is the resistance of the wire making up the coil. We see something interesting happen here: once the motor is moving at full speed, we’ll see that $\varepsilon$ is not very much smaller than $V$, meaning the difference between them is small and the current is therefore also small. At startup, though, the back emf is zero since the motor is still momentarily at rest, and the current is much larger. This is why a motor will briefly draw much more current when starting than it does while running. If you’ve noticed your lights dimming when your refrigerator’s compressor motor starts up, now you know why.
**Mutual Inductance and Self Inductance**

A current in a coil causes a magnetic field, and a changing magnetic field causes a current in a coil. What happens if one coil is near another and the first (primary) coil is connected to a generator? The primary coil will develop a magnetic field, which will of course change with time since the voltage of the generator and therefore the current of the coil oscillates in time. As this changing magnetic field encounters the other coil (not connected to a generator, and called the secondary), it will cause an induced current to flow in it. We see that not only will a magnet moving through a coil induce a current in the coil, a nearby magnetic field which changes in time will also induce a current. This effect is known as **mutual induction**. The emf induced in the secondary coil depends on the change in magnetic flux through the secondary as time goes on. We know that the magnetic flux in the secondary depends on its area, its number of loops, and the magnetic field (produced by the primary). The primary’s magnetic field is proportional to the current through it, so we can say the total flux through the secondary coil $N_s \Phi_s$ is proportional to $I_p$. If the two things are proportional, we can include a constant and set them equal to each other. The constant will be called $M$, for mutual inductance:

$$N_s \Phi_s = M I_p \quad \text{so} \quad M = \frac{N_s \Phi_s}{I_p}$$

which we can rewrite as

$$\varepsilon = -M \frac{\Delta I_p}{\Delta t}$$

The units for mutual inductance are V s/A, which are called **Henries** (H), with 1 H being a rather large value for inductance. Calculating this is much harder than just measuring it unless we have a circuit arrangement of very high symmetry.

A related phenomenon is the emf induced in a coil by its own changing magnetic field, known as **self-induction**. When the current changes sinusoidally in a coil, the changing magnetic field it produces will give us a contribution to the emf in the coil itself. By the same argument we just used to find the relationship between mutual inductance and the change in the current, we find that the formula connecting self-inductance (we’ll use $L$ as the symbol for self-inductance, and the units are still Henries), changing current, and induced emf is

$$\varepsilon = -L \frac{\Delta I}{\Delta t}$$
Notice that we have no subscripts in this formula – there is only one coil here. Placing something inside the coil, like iron, greatly enhances the inductance of the coil. Coils are so associated with this effect that they are generally called **inductors**. We know electrical energy is stored in capacitors; similarly, magnetic energy is stored in inductors. The work needed to push charges through the coil against self-induced emf is just $W = Q \varepsilon$. The formula for $\varepsilon$ is shown above, and it changes as the current changes. In a similar way, we saw that the work done by a spring depends on the displacement of the spring, but the force also depends on the displacement. Ideally, we’d use calculus to find the work, but we can also do it by looking at a graph. We’ll find that

$$W = \frac{1}{2} LI^2$$

Using the fact that $L$ for a solenoid is

$$L = \mu_0 n^2 Al$$

and the magnetic field

$$B = \mu_0 n I$$

and combining them, we get an energy of

$$E = \frac{1}{2 \mu_0} B^2 Al$$

The energy per volume can be found if we divide by the volume of the solenoid $Al$ and we get

$$u = \frac{B^2}{2 \mu_0}$$

**Transformers**

The induced emf in a coil can be put to an interesting use with an AC current. Let’s build two circuits, each with inductors (coils of wire). If the coils overlap (as if they were both wound around the same tube), the changing current in one will induce a changing current in the other. The flux through each loop will be the same, though, and the emf in each circuit will depend on this flux and the number of loops in the solenoid as
\[ \varepsilon_S = -N_S \frac{\Delta \Phi}{\Delta t} \]

where the “s” stands for secondary. The flux will be the same as in the other coil, called the primary, which obeys

\[ \varepsilon_P = -N_P \frac{\Delta \Phi}{\Delta t} \]

Because the change in flux with time will be identical for both inductors, we can see that each coil will have the same value of emf/number of turns or

\[ \frac{\varepsilon_S}{\varepsilon_P} = \frac{N_S}{N_P} \]

This tells us that the voltage in the secondary is related to the voltage in the primary by the ratio of turns in each solenoid. For greater efficiency, rather than threading the coils among each other, each is wound around a side of an iron square. The iron enhances the magnetic field and helps contain the flux so that it goes mainly through the coils. See below

This device is known as a transformer, because it transforms voltages from one level to another level. The primary is connected to the AC source and the power is put to work in the circuit which includes the secondary. If there are more loops in the secondary than the primary, the voltage will be higher in the secondary and the transformer is called a step-up transformer. If there are more loops in the primary, it will be a step-down transformer. Since the transformer is just a piece of iron, it can’t change the overall power coming in, which must match the power going out (except for the small loss of power due to the fact that the transformer isn’t perfect; they’re generally very efficient, but not perfect). If the power is going to stay the same, increasing the voltage must involve decreasing the current by the same factor. A transformer can
change a large, low voltage current into a small, high voltage current or vice versa, but it can’t change the total power. The reason transformers are so important is that the power lost in transmission lines (as in any resistor) goes as the square of the current multiplied by the resistance (which is proportional to the wire’s length). If you need to move lots of power over a long distance, sending a small current at a high voltage minimizes the losses in the wire. For that reason, long-distance power lines are typically in the neighborhood of hundreds of thousands of volts. This is very efficient, but we don’t want this kind of voltage coming into our homes (sparks would be jumping the gap between the two holes in our outlets at this voltage) for safety reasons. Therefore, we step the voltage down at substations and then again on power poles outside our homes. For this reason, AC is vastly superior to DC for our power grid.

**Capacitors and Capacitive Reactance**

In a DC circuit, capacitors and inductors behave strangely; a capacitor will build up charge until its voltage matches the voltage applied to it, and then it will stop, appearing to the rest of the circuit as a broken wire (no current flow). An inductor, however, will act as a straight piece of wire in a DC circuit, doing nothing to the current in steady-state flow (after the circuit has been connected for a fraction of a second and everything has stabilized).

AC circuits are where these elements are valuable to us. In an AC circuit, a capacitor will charge up to a maximum but then discharge as the voltage swings through its peak and towards a minimum. Therefore, positive charge accumulates on one plate of the capacitor while negative charge piles up on the other, and then they rapidly drain off and switch sides as the applied voltage reverses direction. Charge is therefore flowing all the time when a capacitor is connected to an AC circuit.

If we want to talk about the current in the circuit, we can keep our familiar relation between current and voltage (Ohm’s law) if we define something like a resistance for the capacitor. This quantity will be called the **capacitive reactance** and will connect the rms values of the current and voltage as

\[ V_{\text{rms}} = I_{\text{rms}} \times X_C \]

where \( X_C \) is the capacitive reactance and is measured in ohms (as it has to be for the units to work out). The capacitive reactance is connected to the capacitance by

\[ X_C = \frac{1}{2\pi f C} \]

demonstrating a frequency dependence for this quantity. As the frequency of the AC current increases, the capacitive reactance decreases (so the capacitor doesn’t resist the current very much). In the DC limit (where frequency would be 0), the reactance heads off to infinity, signifying the fact that there will be no current flow in an open circuit.
While this explains what happens over one (or more) full cycle(s), it doesn’t cover the instantaneous relationship between current in the circuit containing only a capacitor and voltage across it. Initially, when the voltage is climbing from zero to its maximum, it is easy to put charge on the capacitor’s plates since there is no charge already there. For definiteness, let’s say that there is a left plate and a right plate in the capacitor. The first charges go on easiest, in other words (let’s say positive charges are appearing on the left plate). That means that the rate of flow of charge, or current, is highest while the voltage is at zero. As the voltage increases, more charges are piled on the plates, but the rate at which they’re placed there tapers off since the charges already there fight the addition of any more like charges. By the time the voltage peaks, the capacitor is fully charged and no more charges can be added to the plates, so the current flow is zero. When the voltage turns around and starts heading lower, the charges start bleeding off the capacitor (slowly at first, then faster) so the current starts to increase but opposite to its initial direction. This current hits a negative maximum when the voltage in the circuit is zero and it becomes very easy to add negative charges to the left plate. The cycle continues with voltage and current $\frac{1}{4}$ cycle (90 degrees) out of phase. We say that the current leads the voltage since it will be a maximum when the voltage is just starting out at zero. Using our previous convention that the voltage produced by an AC generator can be written as

$$V = V_0 \sin(2\pi ft)$$

we can add 90 degrees to this to find the instantaneous current in a circuit which is only capacitive and get

$$I = I_0 \cos(2\pi ft)$$

Interestingly, the average power consumed in this circuit (which would be $I \times V$ over a full cycle) will be zero. Energy will just slosh back and forth in the circuit if it only contains a capacitor – for half the cycle, the generator will be charging the capacitor and for the other half, the capacitor will return that energy to the generator.

A real-world circuit will not be purely capacitive (the wires themselves will have a little resistance, if nothing else), so we now have to take that into account. In a purely resistive circuit, current and voltage are always in phase and we can find the power used by just multiplying their rms values together. The fact that current in a capacitive circuit is out of phase with the voltage means that adding the contributions from both resistors and capacitors (and later inductors) will be slightly more complicated.

We’re going to use a special kind of vector which will rotate around the origin of our coordinate system with a frequency $f$. This vector is called a phasor and we’ll have one for the current and another for the voltage (since they aren’t generally reaching their maxima at the same time except in a purely resistive circuit). In the phasor diagram below for a purely resistive circuit, the two phasors representing current and voltage are rotating counterclockwise and making a complete rotation $f$ times per second. The instantaneous values of current or voltage at some time can be found by just looking at the horizontal component of the phasor at that time. You
can see for yourself that the horizontal component will go from zero to a maximum, down to zero, to a minimum, and back up to zero each cycle.

For a circuit that’s purely capacitive, we’ll have the current arrow (blue) 90 degrees in front of the voltage arrow (purple)

Since the arrows are moving counterclockwise, the current arrow (blue) is leading the voltage arrow. While this probably seems like adding a layer of complexity for nothing, we’ll see that it’s very useful when our circuits contain multiple elements (resistors, capacitors, and inductors).

**Inductors and Inductive Reactance**

There’s also a connection between the voltage across an inductor and the current (more precisely, the rate of change of the current in time). In terms of an average, we can define the inductive equivalent of capacitive reactance, called (surprise!) inductive reactance. The current in a purely inductive circuit is connected to the voltage across the inductor by

\[ V_{rms} = I_{rms} X_L \]

where the inductive reactance \( X_L \) depends on frequency and inductance as

\[ X_L = 2\pi f L \]
Notice that as the frequency of the AC current increases, the reactance increases so the current is reduced. The inductor fights high frequency currents. In the case of a DC circuit, where \( f=0 \), the inductive reactance is zero and the inductor acts like a straight wire.

As in the case of the capacitor, the instantaneous values of current and voltage are not in phase. When the current through the circuit is zero, it is changing most rapidly so the voltage produced by the inductor is at its peak (maximum or minimum). Once the current peaks, it drops off relatively slowly, so the change in current with time is small and the voltage produced by the inductor is small. Current and voltage are still out of phase by 90 degrees, but for the inductive circuit, the current lags the voltage rather than leading it. The 90 degree phase difference means that, as in the capacitive case, the purely inductive circuit consumes no power when averaged over a cycle. The phasor diagram for an inductor is therefore opposite to that of the capacitor:

We can explain this by recalling harmonic motion from last semester. The position of an object in simple harmonic motion was described by

\[
x = A \cos \left( 2\pi f t \right)
\]

which is almost exactly the same as our expression for the voltage of a generator in an AC circuit. If we had chosen to start our particle at a position of zero at time zero, we would have instead written

\[
x = A \sin \left( 2\pi f t \right)
\]

just like the generator. Remember that our object moving in simple harmonic motion had a velocity which was 90 degrees out of phase with its position (i.e., if we’re talking about a mass on a spring, its velocity is zero when the position is at its maximum/minimum, but the velocity is highest when the position is zero):

\[
v = 2\pi f A \cos \left( 2\pi f t \right)
\]
Also, the acceleration was a further 90 degrees out of phase with the velocity, so that it hit a negative maximum when the position was at a positive maximum (from $F = ma = -kx$) and both acceleration and position were zero at the same time:

$$a = -4\pi^2 f^2 A \sin(2\pi ft)$$

Remembering that velocity is just the change in position with time and acceleration is just the change in velocity with time, what we see is that each time derivative (the “change in time”) moves us through the cycle of $\sin \rightarrow \cos \rightarrow -\sin \rightarrow -\cos$ and back to $\sin$ and adds another factor of $2\pi f$. How does this connect to circuits?

The relations we’ve seen between voltage and capacitance, resistance, and inductance

$$Q = CV \quad V = IR \quad V = -N \frac{\Delta I}{\Delta t}$$

can make this clear when we realize that

$$I = \frac{\Delta Q}{\Delta t}$$

This is why voltage is in phase with the current in a purely resistive circuit and 90 degrees out of phase (each direction) in purely capacitive or purely inductive circuits. What about a real-world circuit which has all three of these elements in some amount?

**AC Circuits with Inductors, Capacitors, and Resistors**

When a circuit has inductance, capacitance, and resistance, we have to combine them to find their overall effect. While they are all measured in ohms, we can’t add them directly for the same reason that we can’t add the x and y components of a vector. Looking at the phasor diagrams above, we see that the current in a capacitive circuit would be 180 degrees out of phase with the current in a similar inductive circuit. The overall instantaneous voltage is found by taking the difference between voltages associated with the inductive reactance and capacitive reactance (since the voltages connected to them are 180 degrees out of phase, they’re in “opposite directions”) and using the Pythagorean theorem to combine this with the voltage connected to the resistance. The instantaneous voltage is therefore

$$V_0^2 = V_R^2 + (V_L - V_C)^2$$

and we could just as well consider these quantities to be rms instead of instantaneous. We can use the earlier formulae we’ve seen for voltage and current in these circuits and factor out the current, leaving us with (rms values)
\[ V_{rms} = I_{rms} \sqrt{R^2 + (X_L - X_C)^2} \]

We can preserve the form of ohm’s law for these more complicated circuits by defining the \textbf{impedance} of the circuit (measured in ohms and denoted by \(Z\)) as

\[ Z = \sqrt{R^2 + (X_L - X_C)^2} \]

If the inductor, resistor, and capacitor are all in series, the same current has to flow through all of them. The voltage across the inductor will lead the current, the voltage across the capacitor will lag the current, and the voltage across the resistor will be in phase with the current. A picture of this (using phasors) is shown below:

Notice that the current (black arrow) is in the same direction as the voltage across the resistor (red) since they’re in phase. The voltage across the inductor (green) is ahead of the current by 90 degrees, and the voltage across the capacitor (blue) is behind the current by 90 degrees. The overall instantaneous voltage will be found by combining these three voltages as described above, and we’ll get
The net vector (phasor, really) combination of all these voltages is the orange arrow $V_0$ which, in
general, doesn’t point in the same direction as any of the three voltages that made it up. The
angle by which the total instantaneous voltage $V_0$ leads the current $I$ is known as the phase angle
$\phi$ and is shown on the picture above in purple. We find this the same way we’d find the direction
of any other vector – it will be the arctangent of the ratio of the $x$ and $y$ components, specifically

$$Tan\phi = \frac{V_L - V_C}{R} = \frac{X_L - X_C}{R}$$

Part of the importance of this angle is in determining the power used by the circuit. Since $V_0$
points in the same direction as $Z$ (by definition) and $V_R$ points in the same direction as $R$, we can
use the figure above to see that $R = Z \cos \phi$. We can find the average power dissipated by

$$\bar{P} = I_{rms}^2 \quad R = I_{rms}^2 \quad Z \cos \phi = I_{rms} \quad V_{rms} \quad \cos \phi$$

Because of the formula above, $\cos \phi$ is called the power factor for the circuit. For a circuit with
only capacitance and/or inductance, the power factor will be zero because inductors and
capacitors don’t use power (when averaged over a full cycle). Purely resistive circuits have a
power factor of 1.

**Resonance in Electric Circuits**

When a circuit has both capacitance and inductance, energy can flow back and forth between
these two elements. A discharging capacitor produces a changing current that builds up a
magnetic field in an inductor which will produce a voltage acting to charge the capacitor. For a
given combination of inductance and capacitance, there is a particular frequency at which energy
is transferred most efficiently from the generator to the circuit. This resonant frequency $f_0$ is
given by

$$f_0 = \frac{1}{2\pi \sqrt{LC}}$$

At this frequency, the impedance is a minimum so the rms current is a maximum. At this
frequency, what will the inductive and capacitive reactances be?

$$X_C = \frac{1}{2\pi C f_0} = \frac{1}{\sqrt{C/L}} = \sqrt{\frac{L}{C}} \quad \text{and} \quad X_L = 2\pi L f_0 = \sqrt{\frac{L}{C}}$$

If the two are equal, the impedance will just equal the resistance, which is as small as it can be,
making the current as large as it can be.
**Semiconductor Devices**

In between conductors and insulators in electrical conductivity, we find semiconductors such as silicon and germanium. When pure, these materials don’t conduct very well at all, but they can be altered to be better. While the outer 4 electrons in silicon are involved in the binding of an individual Si atom to its neighbors, we can replace some of the Si atoms in this lattice (regular 3-D grid of atoms) with something like phosphorus, which has 5 outer shell electrons. Only 4 of them will participate in binding the atom to the neighboring Si atoms, leaving one electron to roam about freely. If it leaves, the phosphorus atom will have an overall positive charge since it was neutral to begin with and being part of this lattice has let one of its electrons “escape”. Because the roaming charge carriers in this altered semiconductor (also called a doped semiconductor) are electrons (negative), it’s an n-type semiconductor. Instead of using phosphorus, we could also dope the semiconductor with something like boron which has only 3 outer shell electrons, effectively leaving a “hole” which acts like a positive charge. When an electron from a nearby Si atom gets too close to this hole, it can get stuck in it, meaning the hole is now around one of the Si atoms. We can think of this movement of electrons from site to site filling in one hole and leaving another behind them as the movement of a positively charged hole through the material. This is called a p-type semiconductor. Putting these different kinds of semiconductors in contact with one another gives us some interesting electronics.

If we combine one of each of these devices to make something called a p-n junction diode, something interesting will happen at the interface between them. The excess electrons in the n-type semiconductor will cross the junction and fill in the movable holes in the p-type semiconductor. Now the atoms near the junction have 4 electrons in their outer shells and are “happy”, but remember that we had overall neutrality in each device alone before this happened. Now the n-type semiconductor has lost some electrons and therefore has an overall positive charge, while the formerly neutral p-type semiconductor has picked up some electrons and become negatively charged. This means there’s an electric field across the junction, and it tends to inhibit any more exchanges.

We can put this diode in a circuit and we see that if we connect the p-type end to the positive terminal and the n-type end to the negative terminal, a current will flow. Connecting the diode to the battery will cause electrons to flow from the p-type end into the positive end of the battery and electrons will also flow from the negative end of the battery into the n-type end. The electrons in the n-type end will then move towards the junction, while the holes in the p-type end will also move towards the junction. The interesting thing is that both of these events represent a (conventional) current moving in the same direction! This condition is known as forward bias – the diode passes current very easily.

Now, what if we turn this diode around? Connecting the p-type end to the negative terminal and the n-type end to the positive terminal will cause the current to stop. Electrons from the negative end of the battery attract the holes in the p-type region and pull them away from the junction. Similarly, the positive end of the battery will attract electrons in the n-type region and pull them away from the junction. Holes pile up at the connection between the negative terminal and the p-type end, and electrons pile up at the connection between the positive terminal and the n-type end. The current stops (this condition is known as reverse bias). A diode therefore functions as a one-way valve for current. A diode is the first circuit element we’ve seen that is directional –
you don’t have to worry about putting a resistor, capacitor, or inductor into a circuit backwards, but it will cause problems if you do it with a diode. A current vs. voltage plot for a diode looks very different than it does for a resistor. See below

The resistor relates current and voltage very simply, but the (ideal) diode is more complicated. We can use a diode to produce DC (approximately) from AC since it will only pass the current in the positive part of the oscillation. Also, some diodes produce light when the electrons and holes meet at the junction and combine. These are called, unsurprisingly, light emitting diodes (LEDs).