Overview  Having shown in Chapter 5 that the magnetic force must exist, we will now study the various properties of the magnetic field and show how it can be calculated for an arbitrary (steady) current distribution. The Lorentz force gives the total force on a charged particle as \( \mathbf{F} = q \mathbf{E} + q \mathbf{v} \times \mathbf{B} \). The results from the previous chapter give us the form of the magnetic field due to a long straight wire. This form leads to Ampère’s law, which relates the line integral of the magnetic field to the current enclosed by the integration loop. It turns out that Ampère’s law holds for a wire of any shape. When supplemented with a term involving changing electric fields, this law becomes one of Maxwell’s equations (as we will see in Chapter 9). The sources of magnetic fields are currents, in contrast with the sources of electric fields, which are charges; there are no isolated magnetic charges, or monopoles. This statement is another of Maxwell’s equations.

As in the electric case, the magnetic field can be obtained from a potential, but it is now a vector potential; its curl gives the magnetic field. The Biot–Savart law allows us to calculate (in principle) the magnetic field due to any steady current distribution. One distribution that comes up often is that of a solenoid (a coil of wire), whose field is (essentially) constant inside and zero outside. This field is consistent with an Ampère’s-law calculation of the discontinuity of \( \mathbf{B} \) across a sheet of current. By considering various special cases, we derive the Lorentz transformations of the electric and magnetic fields. The electric (or magnetic) field in one frame depends on both the electric and magnetic fields in another frame. The Hall effect arises from the \( q \mathbf{v} \times \mathbf{B} \) part of the
The magnetic field

Lorentz force. This effect allows us, for the first time, to determine the sign of the charge carriers in a current.

6.1 Definition of the magnetic field

A charge that is moving parallel to a current of other charges experiences a force perpendicular to its own velocity. We can see it happening in the deflection of the electron beam in Fig. 5.3. We discovered in Section 5.9 that this is consistent with – indeed, is required by – Coulomb’s law combined with charge invariance and special relativity. And we found that a force perpendicular to the charged particle’s velocity also arises in motion at right angles to the current-carrying wire. For a given current, the magnitude of the force, which we calculated for the particular case in Fig. 5.22(a), is proportional to the product of the particle’s charge $q$ and its speed $v$ in our frame. Just as we defined the electric field $E$ as the vector force on unit charge at rest, so we can define another field $B$ by the velocity-dependent part of the force that acts on a charge in motion. The defining relation was introduced at the beginning of Chapter 5. Let us state it again more carefully.

At some instant $t$ a particle of charge $q$ passes the point $(x, y, z)$ in our frame, moving with velocity $v$. At that moment the force on the particle (its rate of change of momentum) is $F$. The electric field at that time and place is known to be $E$. Then the magnetic field at that time and place is defined as the vector $B$ that satisfies the following vector equation (for any value of $v$):

$$ F = qE + qv \times B $$

This force $F$ is called the Lorentz force. Of course, $F$ here includes only the charge-dependent force and not, for instance, the weight of the particle carrying the charge. A vector $B$ satisfying Eq. (6.1) always exists. Given the values of $E$ and $B$ in some region, we can with Eq. (6.1) predict the force on any particle moving through that region with any velocity. For fields that vary in time and space, Eq. (6.1) is to be understood as a local relation among the instantaneous values of $F$, $E$, $v$, and $B$. Of course, all four of these quantities must be measured in the same inertial frame.

In the case of our “test charge” in the lab frame of Fig. 5.22(a), the electric field $E$ was zero. With the charge $q$ moving in the positive $x$ direction, $v = \hat{x}v$, we found in Eq. (5.28) that the force on it was in the negative $y$ direction, with magnitude $Iqv/2\pi \epsilon_0 rc^2$:

$$ F = -\hat{y} \frac{Iqv}{2\pi \epsilon_0 rc^2}. $$

(6.2)
In this case the magnetic field must be

\[ B = \hat{z} \frac{2I}{2\pi \epsilon_0 rc^2} \]  \hspace{1cm} (6.3)

for then Eq. (6.1) becomes

\[ F = qv \times B = (\hat{x} \times \hat{z})(qv) \left( \frac{I}{2\pi \epsilon_0 rc^2} \right) = -\hat{y} \frac{Iqv}{2\pi \epsilon_0 rc^2}, \]  \hspace{1cm} (6.4)

in agreement with Eq. (6.2).

The relation of \( B \) to \( r \) and to the current \( I \) is shown in Fig. 6.1. Three mutually perpendicular directions are involved: the direction of \( B \) at the point of interest, the direction of a vector \( r \) from the wire to that point, and the direction of current flow in the wire. Here questions of handedness arise for the first time in our study. Having adopted Eq. (6.1) as the definition of \( B \) and agreed on the conventional rule for the vector product, that is, \( \hat{x} \times \hat{y} = \hat{z} \), etc., in coordinates like those of Fig. 6.1, we have determined the direction of \( B \). That relation has a
The magnetic field

handedness, as you can see by imagining a particle that moves along the wire in the direction of the current while circling around the wire in the direction of $B$. Its trail, no matter how you look at it, would form a right-hand helix, like that in Fig. 6.2(a), not a left-hand helix like that in Fig. 6.2(b).

From the $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ relation, we see that another set of three (not necessarily mutually perpendicular) vectors consists of the force $\mathbf{F}$ on the charge $q$, the velocity $\mathbf{v}$ of the charge, and the magnetic field $\mathbf{B}$ at the location of the charge. In Fig. 6.1, $\mathbf{v}$ happens to point along the direction of the wire, and $\mathbf{F}$ along the direction of $r$, but these need not be the directions in general; $\mathbf{F}$ will always be perpendicular to both $\mathbf{v}$ and $\mathbf{B}$, but $\mathbf{v}$ can point in any direction of your choosing, so it need not be perpendicular to $\mathbf{B}$.

Consider an experiment like Oersted’s, as pictured in Fig. 5.2(a). The direction of the current was settled when the wire was connected to the battery. Which way the compass needle points can be stated if we color one end of the needle and call it the head of the arrow. By tradition, long antedating Oersted, the “north-seeking” end of the needle is so designated, and that is the black end of the needle in Fig. 5.2(a). If you compare that picture with Fig. 6.1, you will see that we have defined $\mathbf{B}$ so that it points in the direction of “local magnetic north.” Or, to put it another way, the current arrow and the compass needle in Fig. 5.2(a) define a right-handed helix (see Fig. 6.2), as do the current direction and the vector $\mathbf{B}$ in Fig. 6.1. This is not to say that there is anything intrinsically right-handed about electromagnetism. It is only the self-consistency of our rules and definitions that concerns us here. Let us note, however, that a question of handedness could never arise in electrostatics. In this sense the vector $\mathbf{B}$ differs in character from the vector $\mathbf{E}$. In the same way, a vector representing an angular velocity, in mechanics, differs from a vector representing a linear velocity.

The SI units of $\mathbf{B}$ can be determined from Eq. (6.1). In a magnetic field of unit strength, a charge of one coulomb moving with a velocity of one meter/second perpendicular to the field experiences a force of one newton. The unit of $\mathbf{B}$ so defined is called the tesla:

$$1 \text{ tesla} = 1 \frac{\text{newton}}{\text{coulomb} \cdot \text{meter/second}} = 1 \frac{\text{newton}}{\text{amp} \cdot \text{meter}}. \tag{6.5}$$

In terms of other units, 1 tesla equals $1 \text{ kg C}^{-1} \text{ s}^{-1}$. In SI units, the relation between field and current in Eq. (6.3) is commonly written as

$$\mathbf{B} = \hat{z} \frac{\mu_0 I}{2\pi r} \tag{6.6}$$

We now know that the earth’s magnetic field has reversed many times in geologic history. See Problem 7.19 and the reference there given.
where \( B \) is in teslas, \( I \) is in amps, and \( r \) is in meters. The constant \( \mu_0 \), like the constant \( \epsilon_0 \) we met in electrostatics, is a fundamental constant in the SI unit system. Its value is defined to be exactly

\[
\mu_0 \equiv 4\pi \cdot 10^{-7} \frac{\text{kg m}}{\text{C}^2}
\]  

(6.7)

Of course, if Eq. (6.6) is to agree with Eq. (6.3), we must have

\[
\mu_0 = \frac{1}{\epsilon_0 c^2} \implies c^2 = \frac{1}{\mu_0 \epsilon_0}
\]  

(6.8)

With \( \epsilon_0 \) given in Eq. (1.3), and \( c = 2.998 \cdot 10^8 \text{ m/s} \), you can quickly check that this relation does indeed hold.

**Remark:** Given that we already found the \( B \) field due to a current-carrying wire in Eq. (6.3), you might wonder what the point is of rewriting \( B \) in terms of the newly introduced constant \( \mu_0 \) in Eq. (6.6). The answer is that \( \mu_0 \) is a product of the historical development of magnetism, which should be contrasted with the special-relativistic development we followed in Chapter 5. The connection between electric and magnetic effects was certainly observed long before the formulation of special relativity in 1905. In particular, as we learned in Section 5.1, Oersted discovered in 1820 that a current-carrying wire produces a magnetic field. And \( \mu_0 \) was eventually introduced as the constant of proportionality in Eq. (6.6). (Or, more accurately, \( \mu_0 \) was assigned a given value, and then Eq. (6.6) was used to define the unit of current.) But even with the observed connection between electricity and magnetism, in the mid nineteenth century there was no obvious relation between the \( \mu_0 \) in the expression for \( B \) and the \( \epsilon_0 \) in the expression for \( E \). They were two separate constants in two separate theories. But two developments changed this.

First, in 1861 Maxwell wrote down his set of equations that govern all of electromagnetism. He then used these equations to show that electromagnetic waves exist and travel with speed \( 1/\sqrt{\mu_0 \epsilon_0} \approx 3 \cdot 10^8 \text{ m/s} \). (We’ll study Maxwell’s equations and electromagnetic waves in Chapter 9.) This strongly suggested that light is an electromagnetic wave, a fact that was demonstrated experimentally by Hertz in 1888. Therefore, \( c = 1/\sqrt{\mu_0 \epsilon_0} \), and hence \( \mu_0 = 1/\epsilon_0 c^2 \). This line of reasoning shows that the speed of light \( c \) is determined by the two constants \( \epsilon_0 \) and \( \mu_0 \).

The second development was Einstein’s formulation of the special theory of relativity in 1905. Relativity was the basis of our reasoning in Chapter 5 (the main ingredients of which were length contraction and the relativistic velocity-addition formula), which led to the expression for the magnetic field in Eq. (6.3). A comparison of this equation with the historical expression in Eq. (6.6) yields \( \mu_0 = 1/\epsilon_0 c^2 \). This line of reasoning shows that \( \mu_0 \) is determined by the two constants \( \epsilon_0 \) and \( c \). Of course, having proceeded the way we did in Chapter 5, there is no need to introduce the constant \( \mu_0 \) in Eq. (6.6) when we already have Eq. (6.3). Nevertheless, the convention in SI units is to write \( B \) in the form given in Eq. (6.6). If you wish, you can think of \( \mu_0 \) simply as a convenient shorthand for the more cumbersome expression \( 1/\epsilon_0 c^2 \).

Comparing the previous two paragraphs, it is unclear which derivation of \( \mu_0 = 1/\epsilon_0 c^2 \) is “better.” Is it preferable to take \( \epsilon_0 \) and \( \mu_0 \) as the fundamental constants and then derive, with Maxwell’s help, the value of \( c \), or to take \( \epsilon_0 \) and
c as the fundamental constants and derive, with Einstein’s help, the value of \( \mu_0 \)? The former derivation has the advantage of explaining why \( c \) takes on the value \( 2.998 \cdot 10^8 \) m/s, while the latter has the advantage of explaining how magnetic forces arise from electric forces. In the end, it’s a matter of opinion, based on what information you want to start with.

In Gaussian units, Eq. (6.1) takes the slightly different form

\[
F = qE + \frac{q}{c} \mathbf{v} \times \mathbf{B}. \tag{6.9}
\]

Note that \( \mathbf{B} \) now has the same dimensions as \( \mathbf{E} \), the factor \( \mathbf{v}/c \) being dimensionless. With force \( F \) in dynes and charge \( q \) in esu, the unit of magnetic field strength is the dyne/esu. This unit has a name, the gauss.

There is no special name for the unit dyne/esu when it is used as a unit of electric field strength. It is the same as 1 statvolt/cm, which is the term normally used for unit electric field strength in the Gaussian system. In Gaussian units, the equation analogous to Eq. (6.3) is

\[
\mathbf{B} = \hat{\mathbf{z}} \frac{2I}{rc}. \tag{6.10}
\]

If you repeat the reasoning of Chapter 5, you will see that this \( \mathbf{B} \) is obtained basically by replacing \( \epsilon_0 \) by \( 1/4\pi \) and erasing one of the factors of \( c \) in Eq. (6.3). \( \mathbf{B} \) is in gauss if \( I \) is in esu/s, \( r \) is in cm, and \( c \) is in cm/s.

**Example (Relation between 1 tesla and 1 gauss)** Show that 1 tesla is equivalent to exactly \( 10^4 \) gauss.

**Solution** Consider a setup where a charge of 1 C travels at 1 m/s in a direction perpendicular to a magnetic field with strength 1 tesla. Equations (6.1) and (6.5) tell us that the charge experiences a force of 1 newton. Let us express this fact in terms of the Gaussian force relation in Eq. (6.9). We know that 1 N = 10^5 dyne and 1 C = 3 \cdot 10^9 esu (this “3” isn’t actually a 3; see the discussion below). If we let 1 tesla = \( n \) gauss, with \( n \) to be determined, then the way that Eq. (6.9) describes the given situation is as follows:

\[
10^5 \text{ dyne} = \frac{3 \cdot 10^9 \text{ esu}}{3 \cdot 10^{10} \text{ cm/s} \left(100 \frac{\text{cm}}{\text{s}}\right)} (n \text{ gauss}). \tag{6.11}
\]

Since 1 gauss equals 1 dyne/esu, all the units cancel, and we end up with \( n = 10^4 \), as desired.

Now, the two 3’s in Eq. (6.11) are actually 2.998’s. This is clear in the denominator because the 3 comes from the factor of \( c \). To see why it is the case in the numerator, recall the example in Section 1.4 where we showed that 1 C = 3 \cdot 10^9 esu. If you redo that example and keep things in terms of the constant \( k \) given in Eqs. (1.2) and (1.3), you will find that the number 3 \cdot 10^9 is actually \( \sqrt{10^9}k \) (ignoring the units of \( k \)). But in view of the definition of \( \mu_0 \) in Eq. (6.7), the \( k = 1/4\pi \epsilon_0 \) expression in Eq. (1.3) can be written as \( k = 1/(10^7 \mu_0 \epsilon_0) \). And we know from above that \( 1/\mu_0 \epsilon_0 = c^2 \), hence \( k = 10^{-7}c^2 \) (ignoring the units).
6.1 Definition of the magnetic field

So the number \(3 \cdot 10^9\) is really \(\sqrt{10^9k} = \sqrt{10^2c^2} = 10c\) (ignoring the units), or \(2.998 \cdot 10^9\). Since both of the 3’s in Eq. (6.11) are modified in the same way, the \(n = 10^4\) result is therefore still exact.

Let us use Eqs. (6.1) and (6.6) to calculate the magnetic force between parallel wires carrying current. Let \(r\) be the distance between the wires, and let \(I_1\) and \(I_2\) be the currents which we assume are flowing in the same direction, as shown in Fig. 6.3. The wires are assumed to be infinitely long—a fair assumption in a practical case if they are very long compared with the distance \(r\) between them. We want to predict the force that acts on some finite length \(l\) of one of the wires, due to the entirety of the other wire. The current in wire 1 causes a magnetic field of strength

\[
B_1 = \frac{\mu_0 I_1}{2\pi r}
\]  

(6.12)

at the location of wire 2. Within wire 2 there are \(n_2\) moving charges per meter length of wire, each with charge \(q_2\) and speed \(v_2\). They constitute the current \(I_2\):

\[
I_2 = n_2q_2v_2.
\]  

(6.13)

According to Eq. (6.1), the force on each charge is \(q_2v_2B_1\). The force on each meter length of wire is therefore \(n_2q_2v_2B_1\), or simply \(I_2B_1\). The force on a length \(l\) of wire 2 is then

\[
F = I_2B_1l
\]  

(6.14)

Using the \(B_1\) from Eq. (6.12), this becomes

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

(6.15)

Here \(F\) is in newtons, and \(I_1\) and \(I_2\) are in amps. As the factor \(l/r\) that appears both in Eq. (6.15) and below in Eq. (6.16) is dimensionless, \(l\) and \(r\) could be in any units.

---

2 \(B_1\) is the field inside wire 2, caused by the current in wire 1. When we study magnetic fields inside matter in Chapter 11, we will find that most conductors, including copper and aluminum, but not including iron, have very little influence on a magnetic field. For the present, let us agree to avoid things like iron and other ferromagnetic materials. Then we can safely assume that the magnetic field inside the wire is practically what it would be in vacuum with the same currents flowing.

3 Equation (6.15) has usually been regarded as the primary definition of the ampere in the SI system, \(\mu_0\) being assigned the value \(4\pi \cdot 10^{-7}\). That is to say, one ampere is the current that, flowing in each of two infinitely long parallel wires a distance \(r\) apart, will cause a force of exactly \(2 \cdot 10^{-7}\) newton on a length \(l = r\) of one of the wires. The other SI electrical units are then defined in terms of the ampere. Thus a coulomb is one ampere-second, a volt is one joule/coulomb, and an ohm is one volt/ampere.
The magnetic field

The same exercise carried out in Gaussian units, with Eqs. (6.9) and (6.10), will lead to

\[ F = \frac{2I_1 I_2 l}{c^2 r} . \]  

Equation (6.15) is symmetric in the labels 1 and 2, so the force on an equal length of wire 1 caused by the field of wire 2 must be given by the same formula. We have not bothered to keep track of signs because we know already that currents in the same direction attract one another.

More generally, we can calculate the force on a small piece of current-carrying wire that sits in a magnetic field \( B \). Let the length of the small piece be \( dl \), the linear charge density of the moving charges be \( \lambda \), and the speed of these charges be \( v \). Then the amount of moving charge in the piece is \( dq = \lambda dl \), and the current is \( I = \lambda v \) (in agreement with Eq. (6.13) since \( \lambda = nq \)). Equation (6.1) tells us that the magnetic force on the piece is

\[ dF = dq \times B = (\lambda dl)(v\hat{v}) \times B = (\lambda v)(dl\hat{v}) \times B \]

\[ \implies dF = I \, dl \times B \]

The vector \( dl \) gives both the magnitude and direction of the small piece. The \( F = I_2 B_1 l \) result in Eq. (6.14) is a special case of this result.

**Example (Copper wire)** Let’s apply Eqs. (6.13) and (6.15) to the pair of wires in Fig. 6.4(a). They are copper wires 1 mm in diameter and 5 cm apart. In copper the number of conduction electrons per cubic meter, already mentioned in Chapter 4, is \( 8.45 \times 10^{28} \), so the number of electrons per unit length of wire is \( n = (\pi/4)(10^{-3} \text{ m})^2(8.45 \times 10^{28} \text{ m}^{-3}) = 6.6 \times 10^{22} \text{ m}^{-1} \). Suppose their mean drift velocity \( \bar{v} \) is 0.3 cm/s = 0.003 m/s. (Of course their random speeds are vastly greater.) The current in each wire is then

\[ I = nq\bar{v} = (6.6 \times 10^{22} \text{ m}^{-1})(1.6 \times 10^{-19} \text{ C})(0.003 \text{ m/s}) \approx 32 \text{ C/s} . \]

The attractive force on a 20 cm length of wire is

\[ F = \mu_0 I^2 l = \frac{(4\pi \times 10^{-7} \text{ kg m/C}^2)(32 \text{ C/s})^2(0.2 \text{ m})}{2\pi(0.05 \text{ m})} \approx 8 \times 10^{-4} \text{ N} . \]

This result of \( 8 \times 10^{-4} \text{ N} \) is not an enormous force, but it is easily measurable. Figure 6.4(b) shows how the force on a given length of conductor could be observed.

Recall that the \( \mu_0 \) in Eq. (6.18) can alternatively be written as \( 1/\varepsilon_0 c^2 \). The \( c^2 \) in the denominator reminds us that, as we discovered in Chapter 5, the magnetic force is a relativistic effect, strictly proportional to \( v^2/c^2 \) and traceable to a Lorentz contraction. And with the \( v \) in the above example less than the speed of a healthy ant, it is causing a quite respectable force! The explanation is the immense amount of negative charge the conduction electrons represent, charge that ordinarily is so precisely neutralized by positive charge that we hardly notice it. To appreciate that,
consider the force with which our wires in Fig. 6.4 would repel one another if the charge of the $6.6 \cdot 10^{22}$ electrons per meter were not neutralized at all. As an exercise you can show that the force is just $c^2/v^2$ times the force we calculated above, or roughly $4 \times 10^{15}$ tons per meter of wire. So full of electricity is all matter! If all the electrons in just one raindrop were removed from the earth, the whole earth’s potential would rise by several million volts.

Matter in bulk, from raindrops to planets, is almost exactly neutral. You will find that any piece of it much larger than a molecule contains nearly the same number of electrons as protons. If it didn’t, the resulting electric field would be so strong that the excess charge would be irresistibly blown away. That would happen to electrons in our copper wire even if the excess of negative charge were no more than $10^{-10}$ of the total. A magnetic field, on the other hand, cannot destroy itself in this way. No matter how strong it may be, it exerts no force on a stationary charge. That is why forces that arise from the motion of electric charges can dominate the scene. The second term on the right in Eq. (6.1) can be much larger than the first. Thanks to that second term, an electric motor
The magnetic field

can start your car. In the atomic domain, however, where the coulomb force between pairs of charged particles comes into play, magnetic forces do take second place relative to electrical forces. They are weaker, generally speaking, by just the factor we should expect, the square of the ratio of the particle speed to the speed of light.

Inside atoms we find magnetic fields as large as 10 tesla (or 10^5 gauss). The strongest large-scale fields easily produced in the laboratory are on that order of magnitude too, although fields up to several hundred tesla have been created for short times. In ordinary electrical machinery, electric motors for instance, 1 tesla (or 10^4 gauss) would be more typical. Magnetic resonance imaging (MRI) machines also operate on the order of 1 tesla. A magnet on your refrigerator might have a field of around 10 gauss. The strength of the earth’s magnetic field is a few tenths of a gauss at the earth’s surface, and presumably many times stronger down in the earth’s metallic core where the currents that cause the field are flowing. We see a spectacular display of magnetic fields on and around the sun. A sunspot is an eruption of magnetic field with local intensity of a few thousand gauss. Some other stars have stronger magnetic fields. Strongest of all is the magnetic field at the surface of a neutron star, or pulsar, where in some cases the intensity is believed to reach the hardly conceivable range of 10^10 tesla. On a vaster scale, our galaxy is pervaded by magnetic fields that extend over thousands of light years of interstellar space. The field strength can be deduced from observations in radioastronomy. It is a few microgauss – enough to make the magnetic field a significant factor in the dynamics of the interstellar medium.

6.2 Some properties of the magnetic field

The magnetic field, like the electric field, is a device for describing how charged particles interact with one another. If we say that the magnetic field at the point (4.5, 3.2, 6.0) at 12:00 noon points horizontally in the negative y direction and has a magnitude of 5 gauss, we are making a statement about the acceleration a moving charged particle at that point in space-time would exhibit. The remarkable thing is that a statement of this form, giving simply a vector quantity \( \mathbf{B} \), says all there is to say. With it one can predict uniquely the velocity-dependent part of the force on any charged particle moving with any velocity. It makes unnecessary any further description of the other charged particles that are the sources of

---

4 Nikola Tesla (1856–1943), the inventor and electrical engineer for whom the SI unit was named, invented the alternating-current induction motor and other useful electromagnetic devices. Gauss’s work in magnetism was concerned mainly with the earth’s magnetic field. Perhaps this will help you to remember which is the larger unit. For small magnetic fields, it is generally more convenient to work with gauss than with tesla, even though the gauss technically isn’t part of the SI system of units. This shouldn’t cause any confusion; you can quickly convert to tesla by dividing by (exactly) 10^4. If you’re wary about leaving the familiar ground of SI units, feel free to think of a gauss as a deci-milli-tesla.
the field. In other words, if two quite different systems of moving charges happen to produce the same $E$ and $B$ at a particular point, the behavior of any test particle at the point would be exactly the same in the two systems. It is for this reason that the concept of field, as an intermediary in the interaction of particles, is useful. And it is for this reason that we think of the field as an independent entity.

Is the field more, or less, real than the particles whose interaction, as seen from our present point of view, it was invented to describe? That is a deep question which we would do well to set aside for the time being. People to whom the electric and magnetic fields were vividly real – Faraday and Maxwell, to name two – were led thereby to new insights and great discoveries. Let’s view the magnetic field as concretely as they did and learn some of its properties.

So far we have studied only the magnetic field of a straight wire or filament of steady current. The field direction, we found, is everywhere perpendicular to the plane containing the filament and the point where the field is observed. The magnitude of the field is proportional to $1/r$. The field lines are circles surrounding the filament, as shown in Fig. 6.5. The sense of direction of $B$ is determined by our previously adopted convention about the vector cross-product, by the (arbitrary) decision to write the second term in Eq. (6.1) as $qv \times B$, and by the physical fact that a positive charge moving in the direction of a positive current is attracted to it rather than repelled. These are all consistent if we relate the direction of $B$ to the direction of the current that is its source in the manner shown in Fig. 6.5. Looking in the direction of positive current, we see the $B$ lines curling clockwise. Or you may prefer to remember it as a right-hand-thread relation. Point your right thumb in the direction of the current and your fingers will curl in the direction of $B$.

Let’s look at the line integral of $B$ around a closed path in this field. (Remember that a similar inquiry in the case of the electric field of a point charge led us to a simple and fundamental property of all electrostatic fields, that $\int E \cdot ds = 0$ around a closed path, or equivalently that $\text{curl } E = 0$.) Consider first the path $ABCD$ in Fig. 6.6(a). This lies in a plane perpendicular to the wire; in fact, we need only work in this plane, for $B$ has no component parallel to the wire. The line integral of $B$ around the path shown is zero, for the following reason. Paths $BC$ and $DA$ are perpendicular to $B$ and contribute nothing. Along $AB$, $B$ is stronger in the ratio $r_2/r_1$ than it is along $CD$; but $CD$ is longer than $AB$ by the same factor, for these two arcs subtend the same angle at the wire. So the two arcs give equal and opposite contributions, and the whole integral is zero.

It follows that the line integral is also zero on any path that can be constructed out of radial segments and arcs, such as the path in Fig. 6.6(b). From this it is a short step to conclude that the line integral is zero around any path that does not enclose the wire. To smooth out the corners we would only need to show that the integral around a
The magnetic field

little triangular path vanishes. The same step was involved in the case of the electric field.

A path that does not enclose the wire is one like the path in Fig. 6.6(c), which, if it were made of string, could be pulled free. The line integral around any such path is zero.

Now consider a circular path that encloses the wire, as in Fig. 6.6(d). Here the circumference is $2\pi r$, and the field is $\mu_0 I/2\pi r$ and everywhere parallel to the path, so the value of the line integral around this particular path is $(2\pi r)(\mu_0 I/2\pi r)$, or $\mu_0 I$. We now claim that any path looping once around the wire must give the same value. Consider, for instance, the crooked path $C$ in Fig. 6.6(e). Let us construct the path $C'$ in Fig. 6.6(f) made of a path like $C$ and a circular path, but not enclosing the wire. The line integral around $C'$ must be zero, and therefore the integral around $C$ must be the negative of the integral around the circle, which we have already evaluated as $\mu_0 I$ in magnitude. The sign will depend in an obvious way on the sense of traversal of the path. Our general conclusion is:

$$\int \mathbf{B} \cdot d\mathbf{s} = \mu_0 \times (\text{current enclosed by path})$$

(Ampère’s law).

This is known as Ampère’s law. It is valid for steady currents. In the Gaussian analog of this expression, the $\mu_0$ is replaced with $4\pi/c$, which quickly follows from a comparison of Eqs. (6.6) and (6.10).

Equation (6.19) holds when the path loops the current filament once. Obviously a path that loops it $N$ times, like the one in Fig. 6.6(g), will just give $N$ times as big a result for the line integral.

The magnetic field, as we have emphasized before, depends only on the rate of charge transport, the number of units of charge passing a given point in the circuit, per second. Figure 6.7 shows a circuit with a current of 5 milliamperes. The average velocity of the charge carriers ranges from $10^{-6}$ m/s in one part of the circuit to 0.8 times the speed of light in another. The line integral of $\mathbf{B}$ over a closed path has the same value around every part of this circuit, namely

$$\int \mathbf{B} \cdot d\mathbf{s} = \mu_0 I = \left(4\pi \cdot 10^{-7} \frac{\text{kg m}}{C^2}\right) \left(0.005 \frac{\text{C}}{\text{s}}\right) = 6.3 \cdot 10^{-9} \frac{\text{kg m}}{\text{C s}}.$$

(6.20)

You can check that these units are the same as tesla-meter, which they must be, in view of the left-hand side of this equation.

Figure 6.6.
The line integral of the magnetic field $\mathbf{B}$ over any closed path depends only on the current enclosed.
What we have proved for the case of a long straight filament of current clearly holds, by superposition, for the field of any system of straight filaments. In Fig. 6.8 several wires are carrying currents in different directions. If Eq. (6.19) holds for the magnetic field of one of these wires, it must hold for the total field, which is the vector sum, at every point, of the fields of the individual wires. That is a pretty complicated field. Nevertheless, we can predict the value of the line integral of $\mathbf{B}$ around the closed path in Fig. 6.8 merely by noting which currents the path encircles, and in which sense.

**Example (Magnetic field due to a thick wire)** We know that the magnetic field outside an infinitesimally thin wire points in the tangential direction and has magnitude $B = \mu_0 I / 2\pi r$. But what about a thick wire? Let the wire have radius $R$ and carry current $I$ with uniform current density; the wire may be viewed as the superposition of a large number of thin wires running parallel to each other. Find the field both outside and inside the wire.

**Solution** Consider an Amperian loop (in the spirit of a Gaussian surface) that takes the form of a circle with radius $r$ around the wire. Due to the cylindrical symmetry, $\mathbf{B}$ has the same magnitude at all points on this loop. Also, $\mathbf{B}$ is
The magnetic field

tangential; it has no radial component, due to the symmetric nature of the thin wires being superposed. So the line integral $\int B \cdot ds$ equals $B(2\pi r)$. Ampère’s law then quickly gives $B = \mu_0 I / 2\pi r$. We see that, outside a thick wire, the wire can be treated like a thin wire lying along the axis, as far as the magnetic field is concerned. This is the same result that holds for the electric field of a charged wire.

Now consider a point inside the wire. Since area is proportional to $r^2$, the current contained within a radius $r$ inside the wire is $I_r = I(r^2 / R^2)$. Ampère’s law then gives the magnitude of the (tangential) field at radius $r$ as

$$2\pi r B = \mu_0 I_r \implies B = \frac{\mu_0 I(r^2 / R^2)}{2\pi r} = \frac{\mu_0 I r}{2\pi R^2} \quad (r < R). \quad (6.21)$$

We have been dealing with long straight wires. However, we want to understand the magnetic field of any sort of current distribution – for example, that of a current flowing in a closed loop, a circular ring of current, to take the simplest case. Perhaps we can derive this field too from the fields of the individual moving charge carriers, properly transformed. A ring of current could be a set of electrons moving at constant speed around a circular path. But here that strategy fails us. The trouble is that an electron moving on a circular path is an accelerated charge, whereas the magnetic fields we have rigorously derived are those of charges moving with constant velocity. We shall therefore abandon our program of derivation at this point and state the remarkably simple fact: these more general fields obey exactly the same law, Eq. (6.19). The line integral of $B$ around a bent wire is equal to that around a long straight wire carrying the same current. As this goes beyond anything we have so far deduced, we must look on it here as a postulate confirmed by the experimental tests of its implications.

You may find it unsettling that the validity of Ampère’s law applied to an arbitrarily shaped wire simply has to be accepted, given that we have derived everything up to this point. However, this distinction between acceptance and derivation is illusory. As we will see in Chapter 9, Ampère’s law is a special case of one of Maxwell’s equations. Therefore, accepting Ampère’s law is equivalent to accepting one of Maxwell’s equations. And considering that Maxwell’s equations govern all of electromagnetism (being consistent with countless experimental tests), accepting them is certainly a reasonable thing to do. Likewise, all of our derivations thus far in this book (in particular, the ones in Chapter 5) can be traced back to Coulomb’s law, which is equivalent to Gauss’s law, which in turn is equivalent to another one of Maxwell’s equations. Therefore, accepting Coulomb’s law is equivalent to accepting this other Maxwell equation. In short, everything boils down to Maxwell’s equations sooner or later. Coulomb’s law is no more fundamental than Ampère’s law. We accepted the former long ago, so we shouldn’t be unsettled about accepting the latter now.

To state Ampère’s law in the most general way, we must talk about volume distributions of current. A general steady current distribution is
6.2 Some properties of the magnetic field

6.2 Some properties of the magnetic field described by a current density \( J(x, y, z) \) that varies from place to place but is constant in time. A current in a wire is merely a special case in which \( J \) has a large value within the wire but is zero elsewhere. We discussed volume distributions of current in Chapter 4, where we noted that, for time-independent currents, \( J \) has to satisfy the continuity equation, or conservation-of-charge condition,

\[
\text{div} \ J = 0. \tag{6.22}
\]

Take any closed curve \( C \) in a region where currents are flowing. The total current enclosed by \( C \) is the flux of \( J \) through the surface spanning \( C \), that is, the surface integral \( \int_S J \cdot da \) over this surface \( S \) (see Fig. 6.9). A general statement of the relation in Eq. (6.19) is therefore

\[
\int_C B \cdot ds = \mu_0 \int_S J \cdot da. \tag{6.23}
\]

Let us compare this with Stokes’ theorem, which we developed in Chapter 2:

\[
\int_C F \cdot ds = \int_S (\text{curl} \ F) \cdot da. \tag{6.24}
\]

We see that a statement equivalent to Eq. (6.23) is this:

\[
\text{curl} \ B = \mu_0 J. \tag{6.25}
\]

This is the differential form of Ampère’s law, and it is the simplest and most general statement of the relation between the magnetic field and the moving charges that are its source. As with Eq. (6.19), the Gaussian analog of this expression has the \( \mu_0 \) replaced by \( 4\pi/\epsilon \). Note that the form of \( J \) in Eq. (6.25) guarantees that Eq. (6.22) is satisfied, because the divergence of the curl is always zero (see Exercise 2.78).

**Example (Curl of \( B \) for a thick wire)** For the above “thick wire” example, verify that \( \text{curl} \ B = \mu_0 J \) both inside and outside the wire.

**Solution** We can use the expression for the curl in cylindrical coordinates given in Eq. (F.2) in Appendix F. The only nonzero derivative in the expression is \( \partial (rA_\theta) / \partial r \), so outside the wire we have

\[
\text{curl} \ B = \hat{z} \frac{1}{r} \frac{\partial (rB_\theta)}{\partial r} = \hat{z} \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\mu_0 I}{2\pi r} \right) = 0, \tag{6.26}
\]

which is correct because there is zero current density outside the wire. For the present purposes, the only relevant fact about the external field is that it is proportional to \( 1/r \).

Inside the wire we have

\[
\text{curl} \ B = \hat{z} \frac{1}{r} \frac{\partial (rB_\theta)}{\partial r} = \hat{z} \frac{1}{r} \frac{\partial}{\partial r} \left( \frac{\mu_0 I}{2\pi R^2} \right) = \hat{z} \mu_0 \frac{I}{\pi R^2} = \mu_0 \hat{z} J = \mu_0 J. \tag{6.27}
\]

as desired.
Equation (6.25) by itself is not enough to determine $B(x, y, z)$, given $J(x, y, z)$, because many different vector fields could have the same curl. We need to complete it with another condition. We had better think about the divergence of $B$. Going back to the magnetic field of a single straight wire, we observe that the divergence of that field is zero. You can’t draw a little box anywhere, even one enclosing the wire, that will have a net outward or inward flux. It is enough to note that the boxes $V_1$ and $V_2$ in Fig. 6.10 have no net flux and can shrink to zero without developing any. (The $1/r$ dependence of $B$ isn’t important here. All that matters is that $B$ points in the tangential direction and that its magnitude is independent of $\theta$.) For this field then, $\text{div} \, B = 0$, and hence also for all superpositions of such fields. Again we postulate that the principle can be extended to the field of any distribution of currents, so that a companion to Eq. (6.22) is the condition

$$\text{div} \, B = 0$$  \hspace{1cm} (6.28)

You can quickly check that this relation holds for the wire in the above example, both inside and outside, by using the cylindrical-coordinate expression for the divergence given in Eq. (F.2) in Appendix F; the only nonzero component of $B$ is $B_\theta$, but $\partial B_\theta / \partial \theta = 0$.

We are concerned with fields whose sources lie within some finite region. We won’t consider sources that are infinitely remote and infinitely strong. Under these conditions, $B$ goes to zero at infinity. With this proviso, we have the following theorem.

**Theorem 6.1**  *Assuming that $B$ vanishes at infinity, Eqs. (6.25) and (6.28) together determine $B$ uniquely if $J$ is given.*

**Proof**  Suppose both equations are satisfied by two different fields $B_1$ and $B_2$. Then their difference, the vector field $D = B_1 - B_2$, is a field with zero divergence and zero curl everywhere. What could it be like? Having zero curl, it must be the gradient\(^5\) of some potential function $f(x, y, z)$, that is, $D = \nabla f$. But $\nabla \cdot D = 0$, too, so $\nabla \cdot \nabla f$ or $\nabla^2 f = 0$ everywhere. Over a sufficiently remote enclosing boundary, $f$ must take on some constant value $f_0$, because $B_1$ and $B_2$ (and hence $D$) are essentially zero very far away from the sources. Since $f$ satisfies Laplace’s equation everywhere inside that boundary, it cannot have a maximum or a minimum anywhere in that region (see Section 2.12), and so it must have the value $f_0$ everywhere. Hence $D = \nabla f = 0$, and $B_1 = B_2$. \(\square\)

The fact that a vector field is uniquely determined by its curl and divergence (assuming that it goes to zero at infinity) is known as the

---

\(^5\) This follows from our work in Chapter 2. If $\text{curl} \, D = 0$, then the line integral of $D$ around any closed path is zero. This implies that we can uniquely define a potential function $f$ as the line integral of $D$ from an arbitrary reference point. It then follows that $D$ is the gradient of $f$. 
Helmholtz theorem. We proved this theorem in the special case where the divergence is zero.

In the case of the electrostatic field, the counterparts of Eqs. (6.25) and (6.28) were

\[ \text{curl } E = 0 \quad \text{and} \quad \text{div } E = \frac{\rho}{\epsilon_0} \]  \hspace{1cm} (6.29)

In the case of the electric field, however, we could begin with Coulomb’s law, which expressed directly the contribution of each charge to the electric field at any point. Here we shall have to work our way back to some relation of that type. We shall do so by means of a potential function.

### 6.3 Vector potential

We found that the scalar potential function \( \phi(x, y, z) \) gave us a simple way to calculate the electrostatic field of a charge distribution. If there is some charge distribution \( \rho(x, y, z) \), the potential at any point \( (x_1, y_1, z_1) \) is given by the volume integral

\[
\phi(x_1, y_1, z_1) = \frac{1}{4\pi \epsilon_0} \int \frac{\rho(x_2, y_2, z_2) \, dv_2}{r_{12}}. \quad \hspace{1cm} (6.30)
\]

The integration is extended over the whole charge distribution, and \( r_{12} \) is the magnitude of the distance from \( (x_2, y_2, z_2) \) to \( (x_1, y_1, z_1) \). The electric field \( E \) is obtained as the negative of the gradient of \( \phi \):

\[
E = -\text{grad } \phi. \quad \hspace{1cm} (6.31)
\]

The same trick won’t work for the magnetic field, because of the essentially different character of \( B \). The curl of \( B \) is not necessarily zero, so \( B \) can’t, in general, be the gradient of a scalar potential. However, we know another kind of vector derivative, the curl. It turns out that we can usefully represent \( B \), not as the gradient of a scalar function, but as the curl of a vector function, like this:

\[
B = \text{curl } A \quad \hspace{1cm} (6.32)
\]

By obvious analogy, we call \( A \) the vector potential. It is not obvious, at this point, why this tactic is helpful. That will have to emerge as we proceed. It is encouraging that Eq. (6.28) is automatically satisfied, since \( \text{div curl } A = 0 \), for any \( A \). Or, to put it another way, the fact that \( \text{div } B = 0 \) presents us with the opportunity to represent \( B \) as the curl of another vector function.

---

6 The student may wonder why we couldn’t have started from some equivalent of Coulomb’s law for the interaction of currents. The answer is that a piece of a current filament, unlike an electric charge, is not an independent object that can be physically isolated. You cannot perform an experiment to determine the field from part of a circuit; if the rest of the circuit isn’t there, the current can’t be steady without violating the continuity condition.
The magnetic field

Example (Vector potential for a wire)  As an example of a vector potential, consider a long straight wire carrying a current $I$. In Fig. 6.11 we see the current coming toward us out of the page, flowing along the positive $z$ axis. Outside the wire, what is the vector potential $\mathbf{A}$?

Solution  We know what the magnetic field of the straight wire looks like. The field lines are circles, as sketched already in Fig. 6.5. A few are shown in Fig. 6.11. The magnitude of $\mathbf{B}$ is $\mu_0 I / 2\pi r$. Using a unit vector $\hat{\theta}$ in the tangential direction, we can write the vector $\mathbf{B}$ as

$$
\mathbf{B} = \frac{\mu_0 I}{2\pi r} \hat{\theta}.
$$

(6.33)

We want to find a vector field $\mathbf{A}$ whose curl equals this $\mathbf{B}$. Equation (F.2) in Appendix F gives the expression for the curl in cylindrical coordinates. In view of Eq. (6.33), we are concerned only with the $\hat{\theta}$ component of the curl expression, which is $(\partial A_r / \partial z - \partial A_z / \partial r) \hat{\theta}$. Due to the symmetry along the $z$ axis, we can't have any $z$ dependence, so we are left with only the $-\left(\partial A_z / \partial r\right) \hat{\theta}$ term. Equating this with the $\mathbf{B}$ in Eq. (6.33) gives

$$
\nabla \times \mathbf{A} = \mathbf{B} \implies -\frac{\partial A_z}{\partial r} = \frac{\mu_0 I}{2\pi r} \implies \mathbf{A} = -z \frac{\mu_0 I}{2\pi} \ln r.
$$

(6.34)

This last step can formally be performed by separating variables and integrating. But there is no great need to do this, because we know that the integral of $1/r$ is $\ln r$. The task of Problem 6.4 is to use Cartesian coordinates to verify that the above $\mathbf{A}$ has the correct curl. See also Problem 6.5.

Of course, the $\mathbf{A}$ in Eq. (6.34) is not the only function that could serve as the vector potential for this particular $\mathbf{B}$. To this $\mathbf{A}$ could be added any vector function with zero curl. The above result holds for the space outside the wire. Inside the wire, $\mathbf{B}$ is different, so $\mathbf{A}$ must be different also. It is not hard to find the appropriate vector potential function for the interior of a solid round wire; see Exercise 6.43.

Our job now is to discover a general method of calculating $\mathbf{A}$, when the current distribution $\mathbf{J}$ is given, so that Eq. (6.32) will indeed yield the correct magnetic field. In view of Eq. (6.25), the relation between $\mathbf{J}$ and $\mathbf{A}$ is

$$
curl (\curl \mathbf{A}) = \mu_0 \mathbf{J}.
$$

(6.35)

Equation (6.35), being a vector equation, is really three equations. We shall work out one of them, say the $x$-component equation. The $x$ component of $\curl \mathbf{B}$ is $\partial B_z / \partial y - \partial B_y / \partial z$. The $z$ and $y$ components of $\mathbf{B}$ are, respectively,

$$
B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}, \quad B_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}.
$$

(6.36)

Thus the $x$-component part of Eq. (6.35) reads

$$
\frac{\partial}{\partial y} \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) - \frac{\partial}{\partial z} \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) = \mu_0 J_x.
$$

(6.37)
We assume our functions are such that the order of partial differentiation can be interchanged. Taking advantage of that and rearranging a little, we can write Eq. (6.37) in the following way:

\[- \frac{\partial^2 A_x}{\partial y^2} - \frac{\partial^2 A_x}{\partial z^2} + \frac{\partial}{\partial x} \left( \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) = \mu_0 J_x. \quad (6.38)\]

To make the thing more symmetrical, let’s add and subtract the same term, \(\frac{\partial^2 A_x}{\partial x^2}\), on the left:

\[- \frac{\partial^2 A_x}{\partial x^2} - \frac{\partial^2 A_x}{\partial y^2} - \frac{\partial^2 A_x}{\partial z^2} + \frac{\partial}{\partial x} \left( \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) = \mu_0 J_x. \quad (6.39)\]

We can now recognize the first three terms as the negative of the Laplacian of \(A_x\). The quantity in parentheses is the divergence of \(A\). Now, we have a certain latitude in the construction of \(A\). All we care about is its curl; its divergence can be anything we like. Let us require that

\[\text{div } A = 0. \quad (6.40)\]

In other words, among the various functions that might satisfy our requirement that \(\text{curl } A = B\), let us consider as candidates only those that also have zero divergence. To see why we are free to do this, suppose we had an \(A\) such that \(\text{curl } A = B\), but \(\text{div } A = f(x, y, z) \neq 0\). We claim that, for any function \(f\), we can always find a field \(F\) such that \(\text{curl } F = 0\) and \(\text{div } F = -f\). If this claim is true, then we can replace \(A\) with the new field \(A + F\). This field has its curl still equal to the desired value of \(B\), while its divergence is now equal to the desired value of zero. And the claim is indeed true, because if we treat \(-f\) like the charge density \(\rho\) that generates an electrostatic field, we obviously can find a field \(F\), the analog of the electrostatic \(\mathbf{E}\), such that \(\text{curl } F = 0\) and \(\text{div } F = -f\); the prescription is given in Fig. 2.29(a), without the \(\epsilon_0\).

With \(\text{div } A = 0\), the quantity in parentheses in Eq. (6.39) drops away, and we are left simply with

\[\frac{\partial^2 A_x}{\partial x^2} + \frac{\partial^2 A_x}{\partial y^2} + \frac{\partial^2 A_x}{\partial z^2} = -\mu_0 J_x, \quad (6.41)\]

where \(J_x\) is a known scalar function of \(x, y, z\). Let us compare Eq. (6.41) with Poisson’s equation, Eq. (2.73), which reads

\[\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = -\frac{\rho}{\epsilon_0}. \quad (6.42)\]

The two equations are identical in form. We already know how to find a solution to Eq. (6.42). The volume integral in Eq. (6.30) is the

---

7 This equation is the \(x\) component of the vector identity, \(\nabla \times (\nabla \times \mathbf{A}) = -\nabla^2 \mathbf{A} + \nabla (\nabla \cdot \mathbf{A})\). So in effect, what we’ve done here is prove this identity. Of course, we could have just invoked this identity and applied it to Eq. (6.35), skipping all of the intermediate steps. But it’s helpful to see the proof.
prescription. Therefore a solution to Eq. (6.41) must be given by Eq. (6.30), with \( \rho/\epsilon_0 \) replaced by \( \mu_0 J \):

\[
A_x(x_1, y_1, z_1) = \frac{\mu_0}{4\pi} \int \frac{J_x(x_2, y_2, z_2) \, dv_2}{r_{12}}.
\]  

(6.43)

The other components must satisfy similar formulas. They can all be combined neatly in one vector formula:

\[
A(x_1, y_1, z_1) = \frac{\mu_0}{4\pi} \int \frac{J(x_2, y_2, z_2) \, dv_2}{r_{12}}.
\]  

(6.44)

In more compact notation we have

\[
A = \frac{\mu_0}{4\pi} \int \frac{J \, dv}{r} \quad \text{or} \quad dA = \frac{\mu_0 J \, dv}{4\pi r}.
\]  

(6.45)

There is only one snag. We stipulated that \( \text{div} \, A = 0 \), in order to get Eq. (6.41). If the divergence of the \( A \) in Eq. (6.44) isn’t zero, then although this \( A \) will satisfy Eq. (6.41), it won’t satisfy Eq. (6.39). That is, it won’t satisfy Eq. (6.35). Fortunately, it turns out that the \( A \) in Eq. (6.44) does indeed satisfy \( \text{div} \, A = 0 \), provided that the current is steady (that is, \( \nabla \cdot J = 0 \)), which is the type of situation we are concerned with. You can prove this in Problem 6.6. The proof isn’t important for what we will be doing; we include it only for completeness.

Incidentally, the \( A \) for the example above could not have been obtained by Eq. (6.44). The integral would diverge owing to the infinite extent of the wire. This may remind you of the difficulty we encountered in Chapter 2 in setting up a scalar potential for the electric field of a charged wire. Indeed the two problems are very closely related, as we should expect from their identical geometry and the similarity of Eqs. (6.44) and (6.30). We found in Eq. (2.22) that a suitable scalar potential for the line charge problem is \(- (\lambda/2\pi \epsilon_0) \ln r + C\), where \( C \) is an arbitrary constant. This assigns zero potential to some arbitrary point that is neither on the wire nor an infinite distance away. Both that scalar potential and the vector potential of Eq. (6.34) are singular at the origin and at infinity. However, see Problem 6.5 for a way to get around this issue. For an interesting discussion of the vector potential, including its interpretation as “electromagnetic momentum,” see Semon and Taylor (1996).

### 6.4 Field of any current-carrying wire

Figure 6.12 shows a loop of wire carrying current \( I \). The vector potential \( A \) at the point \((x_1, y_1, z_1)\) is given according to Eq. (6.44) by the integral over the loop. For current confined to a thin wire we may take as the volume element \( dv_2 \) a short section of the wire of length \( dl \). The current density \( J \) is \( I/a \), where \( a \) is the cross-sectional area and \( dv_2 = a \, dl \).
Hence \( J \, dv_2 = I \, dl \), and if we make the vector \( dl \) point in the direction of positive current, we can simply replace \( J \, dv_2 \) by \( I \, dl \). Thus for a thin wire or filament, we can write Eq. (6.44) as a line integral over the circuit:

\[
A = \frac{\mu_0 I}{4\pi} \int \frac{dl}{r_{12}}.
\] (6.46)

To calculate \( A \) everywhere and then find \( B \) by taking the curl of \( A \) might be a long job. It will be more useful to isolate one contribution to the line integral for \( A \), the contribution from the segment of wire at the origin, where the current happens to be flowing in the \( x \) direction (Fig. 6.13). We shall denote the length of this segment by \( dl \). Let \( dA \) be the contribution of this part of the integral to \( A \). Then at the general point \((x, y, z)\), the vector \( dA \), which points in the positive \( x \) direction, is

\[
dA = \hat{x} \frac{\mu_0 l}{4\pi} \frac{dl}{\sqrt{x^2 + y^2 + z^2}}.
\] (6.47)

Let us denote the corresponding part of \( B \) by \( dB \). If we consider now a point \((x, y, 0)\) in the \( xy \) plane, then, when taking the curl of \( dA \) to obtain \( dB \), we find that only one term among the various derivatives survives:

\[
dB = \text{curl} (dA) = \hat{z} \left( -\frac{\partial A_x}{\partial y} \right) = \hat{z} \frac{\mu_0 l}{4\pi} \frac{y \, dl}{(x^2 + y^2)^{3/2}} = \hat{z} \frac{\mu_0 l \sin \phi \, dl}{4\pi} r^2,
\] (6.48)

where \( \phi \) is indicated in Fig. 6.13. You should convince yourself why the symmetry of the \( dA \) in Eq. (6.47) with respect to the \( xy \) plane implies that \( \text{curl} (dA) \) must be perpendicular to the \( xy \) plane.

With this result we can free ourselves at once from a particular coordinate system. Obviously all that matters is the relative orientation of the
The magnetic field element $dl$ and the radius vector $r$ from that element to the point where the field $B$ is to be found. The contribution to $B$ from any short segment of wire $dl$ can be taken to be a vector perpendicular to the plane containing $dl$ and $r$, of magnitude $(\mu_0 I / 4\pi) \sin \phi \, dl / r^2$, where $\phi$ is the angle between $dl$ and $r$. This can be written compactly using the cross-product and is illustrated in Fig. 6.14. In terms of either $\hat{r}$ or $r$, we have

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{d\mathbf{l} \times \hat{r}}{r^2}$$

or

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{d\mathbf{l} \times r}{r^3} \quad \text{(Biot–Savart law).}$$

(6.49)

If you are familiar with the rules of the vector calculus, you can take a shortcut from Eq. (6.46) to Eq. (6.49) without making reference to a coordinate system. Writing $d\mathbf{B} = \nabla \times d\mathbf{A}$, with $d\mathbf{A} = (\mu_0 I / 4\pi) \, dl / r$, we can use the vector identity $\nabla \times (f \mathbf{F}) = f \nabla \times \mathbf{F} + \nabla f \times \mathbf{F}$ to obtain

$$d\mathbf{B} = \nabla \times \frac{\mu_0 I}{4\pi} \frac{d\mathbf{l}}{r} = \frac{\mu_0 I}{4\pi} \left( \frac{1}{r} \nabla \times d\mathbf{l} + \nabla \left( \frac{1}{r} \right) \times d\mathbf{l} \right).$$

(6.50)

But $dl$ is a constant, so the first term on the right-hand side is zero. And recall that $\nabla (1/r) = -\hat{r} / r^2$ (as in going from the Coulomb potential to the Coulomb field). Thus

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \left( -\frac{\hat{r}}{r^2} \right) \times d\mathbf{l} = \frac{\mu_0 I}{4\pi} \frac{d\mathbf{l} \times \hat{r}}{r^2}.$$  

(6.51)

Historically, Eq. (6.49) is known as the Biot–Savart law. The meaning of Eq. (6.49) is that, if $B$ is computed by integrating over the complete circuit, taking the contribution from each element to be given by this formula, the resulting $B$ will be correct. As we remarked in Footnote 6, the contribution of part of a circuit is not physically identifiable. In fact, Eq. (6.49) is not the only formula that could be used to get a correct result for $B$ – to it could be added any function that would give zero when integrated around a closed path.

The Biot–Savart law is valid for steady currents (or for sufficiently slowly changing currents). There is no restriction on the speed of the charges that make up the steady current, provided that they are essentially continuously distributed. The speeds can be relativistic, and the Biot–Savart law still works fine. If the current isn’t steady (and is changing rapidly enough) then, although the Biot–Savart law isn’t valid, a somewhat similar law that involves the so-called “retarded time” is valid. We won’t get into that here, but see Problem 6.28 if you want to get a sense of what the retarded time is all about.

---

Footnote 6: There are actually a few other conditions under which the law is valid, but we won’t worry about those here. See Griffiths and Heald (1991) for everything you might want to know about the conditions under which the various laws of electricity and magnetism are valid.
We seem to have discarded the vector potential as soon as it performed one essential service for us. Indeed, it is often easier, as a practical matter, to calculate the field of a current system directly, now that we have Eq. (6.49), than to find the vector potential first. We shall practice on some examples in Section 6.5. However, the vector potential is important for deeper reasons. For one thing, it has revealed to us a striking parallel between the relation of the electrostatic field $E$ to its sources (static electric charges) and the relation of the magnetostatic field $B$ to its sources (steady electric currents; that’s what magnetostatic means). Its greatest usefulness becomes evident in more advanced topics, such as electromagnetic radiation and other time-varying fields.

### 6.5 Fields of rings and coils

We will now do two examples where we use Eq. (6.49) to calculate a magnetic field. The second example will build on the result of the first.

---

**Example (Circular ring)** A current filament in the form of a circular ring of radius $b$ is shown in Fig. 6.15(a). We could predict without any calculation that the magnetic field of this source must look something like Fig. 6.15(b), where we have sketched some field lines in a plane through the axis of symmetry. The field as a whole must be rotationally symmetrical about this axis, the $z$ axis in Fig. 6.15(a), and the field lines themselves (ignoring their direction) must be symmetrical with respect to the plane of the loop, the $xy$ plane. Very close to the filament the field will resemble that near a long straight wire, since the distant parts of the ring are there relatively unimportant.

It is easy to calculate the field on the axis, using Eq. (6.49). Each element of the ring of length $dl$ contributes a $dB$ perpendicular to $r$. We need only include the $z$ component of $dB$, for we know the total field on the axis must point in the $z$ direction. This brings in a factor of $\cos \theta$, so we obtain

$$dB_z = \frac{\mu_0 I}{4\pi} \frac{dl \cos \theta}{r^2} = \frac{\mu_0 I}{4\pi} \frac{dl b}{r^2} r.$$  \hspace{1cm} (6.52)

Integrating over the whole ring, we have simply $\int dl = 2\pi b$, so the field on the axis at any point $z$ is

$$B_z = \frac{\mu_0 I}{4\pi} \frac{2\pi b^2}{r^3} = \frac{\mu_0 I b^2}{2(b^2 + z^2)^{3/2}} \quad \text{(field on axis)}. \hspace{1cm} (6.53)$$

At the center of the ring, $z = 0$, the magnitude of the field is

$$B_z = \frac{\mu_0 I}{2b} \quad \text{(field at center)}. \hspace{1cm} (6.54)$$

Note that the field points in the same direction (upward) everywhere along the $z$ axis.

---

9 The main reason for this is that if we want to use $A$ to calculate $B$ at a given point, we need to know what $A$ is at nearby points too. That is, we need to know $A$ as a function of the coordinates so that we can calculate the derivatives in the curl. On the other hand, if we calculate $B$ via Eq. (6.49), we simply need to find $B$ at the one given point.

---

**Figure 6.15.**

The magnetic field of a ring of current.

(a) Calculation of the field on the axis. (b) Some field lines.
The magnetic field

Example (Solenoid) The cylindrical coil of wire shown in Fig. 6.16(a) is usually called a solenoid. We assume the wire is closely and evenly spaced so that the number of turns in the winding, per meter length along the cylinder, is a constant, $n$. Now, the current path is actually helical, but if the turns are many and closely spaced, we can ignore this and regard the whole solenoid as equivalent to a stack of current rings. Then we can use Eq. (6.53) as a basis for calculating the field at any point, such as the point $z$, on the axis of the coil.

Consider the contribution from the current rings included between radii from the point $z$ that make angles $\theta$ and $\theta + d\theta$ with the axis. The length of this segment of the solenoid, indicated in Fig. 6.16(b), is $rd\theta/\sin\theta$ (because $rd\theta$ is the tilted angular span of the segment, and the factor of $1/\sin\theta$ gives the vertical span). It is therefore equivalent to a ring carrying a current $dI = In(rd\theta/\sin\theta)$.

Since $r = b/\sin\theta$, Eq. (6.53) gives, for the contribution of this ring to the axial field:

$$dB_z = \frac{\mu_0 b^2}{2r^3} \left( \frac{nI r d\theta}{\sin\theta} \right) = \frac{\mu_0 b^2}{2r^3} \sin\theta d\theta. \quad (6.55)$$

Carrying out the integration between the limits $\theta_1$ and $\theta_2$, gives

$$B_z = \frac{\mu_0 n I}{2} \int_{\theta_1}^{\theta_2} \sin\theta d\theta = \frac{\mu_0 n I}{2} (\cos\theta_1 - \cos\theta_2). \quad (6.56)$$

We have used Eq. (6.56) to make a graph, in Fig. 6.17, of the field strength on the axis of a coil, the length of which is four times its diameter. The ordinate is the field strength $B_z$ relative to the field strength in a coil of infinite length with the same number of turns per meter and the same currents in each turn. For the infinite coil, $\theta_1 = 0$ and $\theta_2 = \pi$, so

$$B_z = \mu_0 n I \quad \text{(infinitely long solenoid).} \quad (6.57)$$

At the center of the “four-to-one” coil the field is very nearly as large as this, and it stays pretty nearly constant until we approach one of the ends. Equation (6.57) actually holds for all points inside an infinite solenoid, not just for points on the axis; see Problem 6.19.

Figure 6.18 shows the magnetic field lines in and around a coil of these proportions. Note that some field lines actually penetrate the winding. The cylindrical sheath of current is a surface of discontinuity for the magnetic field. Of course, if we were to examine the field very closely in the neighborhood of the wires, we would not find any infinitely abrupt kinks, but we would find a very complicated, ripply pattern around and through the individual wires.

It is quite possible to make a long solenoid with a single turn of a thin ribbonlike conductor, as in Fig. 6.19. To this, our calculation and the diagram in Fig. 6.18 apply exactly, the quantity $nI$ being merely replaced by the current per meter flowing in the sheet. Now the change in direction of a field line that penetrates the wall occurs entirely within the thickness of the sheet, as suggested in the inset in Fig. 6.19.

In calculating the field of the solenoid in Fig. 6.16, we treated it as a stack of rings, ignoring the longitudinal current that must exist in
any coil in which the current enters at one end and leaves at the other. Let us see how the field is modified if that is taken into account. The helical coil in Fig. 6.20(c) is equivalent, so far as the field external to the solenoid is concerned, to the superposition of the stack of current rings in Fig. 6.20(a) and a single axial conductor in Fig. 6.20(b). Adding the field of the latter, $B'$, to the field $B$ of the former, we get the external field of the coil. It has a helical twist. Some field lines have been sketched in Fig. 6.20(c). As for the field inside the solenoid, the longitudinal current $I$ flows, in effect, on the cylinder itself. Such a current distribution, a uniform hollow tube of current, produces zero field inside the cylinder
**Figure 6.19.**
A solenoid formed by a single cylindrical conducting sheet. Inset shows how the field lines change direction inside the current-carrying conductor.

**Figure 6.20.**
The helical coil (c) is equivalent to a stack of circular rings, each carrying current \( I \) and shown in (a), plus a current \( I \) parallel to the axis of the coil as shown in (b). A path around the coil encloses the current \( I \), the field of which, \( B' \), must be added to the field \( B \) of the rings to form the external field of the helical coil.

(due to Eq. (6.19) and the fact that a circular path inside the tube encloses no current), leaving unmodified the interior field we calculated before. If you follow a looping field line from inside to outside to inside again, you will discover that it does not close on itself. Field lines generally don’t. You might find it interesting to figure out how this picture would
be changed if the wire that leads the current $I$ away from the coil were brought down along the axis of the coil to emerge at the bottom.

### 6.6 Change in $B$ at a current sheet

In the setup of Fig. 6.19 we had a solenoid constructed from a single curved sheet of current. Let’s look at something even simpler, a flat, unbounded current sheet. You may think of this as a sheet of copper of uniform thickness in which a current flows with constant density and direction everywhere within the metal. In order to refer to directions, let us locate the sheet in the $xz$ plane and let the current flow in the $x$ direction. As the sheet is supposed to be of infinite extent with no edges, it is hard to draw a picture of it! We show a broken-out fragment of the sheet in Fig. 6.21, in order to have something to draw; you must imagine the rest of it extending over the whole plane. The thickness of the sheet will not be very important, finally, but we may suppose that it has some definite thickness $d$.

If the current density inside the metal is $J$ in C s$^{-1}$ m$^{-2}$, then every length $l$ of height, in the $z$ direction, includes a ribbon of current amounting to $J(ld)$ in C/s. We call $Jd$ the surface current density or sheet current density and use the symbol $\mathcal{J}$ to distinguish it from the volume current density $J$. The units$^{10}$ of $\mathcal{J}$ are amps/meter; multiplying $\mathcal{J}$ by the length $l$ of a line segment (perpendicular to the current flow) on the surface gives the current crossing that segment. If we are not concerned with what goes on inside the sheet itself, $\mathcal{J}$ is a useful quantity. It is $\mathcal{J}$ that determines the change in the magnetic field from one side of the sheet to the other, as we shall see.

The field in Fig. 6.21 is not merely that due to the sheet alone. Some other field in the $z$ direction is present, from another source. The total field, including the effect of the current sheet, is represented by the $B$ vectors drawn in front of and behind the sheet.

Consider the line integral of $B$ around the rectangle 12341 in Fig. 6.21. One of the long sides is in front of the surface, the other behind it, with the short sides piercing the sheet. Let $B_z^+$ denote the $z$ component of the magnetic field immediately in front of the sheet, $B_z^-$ the $z$ component of the field immediately behind the sheet. We mean here the field of all sources that may be around, including the sheet itself. The line integral of $B$ around the long rectangle is simply $l(B_z^+ - B_z^-)$.

$^{10}$ The terms “surface” current density and “volume” current density indicate the dimension of the space in which the current flows. Since the units of $\mathcal{J}$ and $J$ are A/m and A/m$^2$, respectively, we are using these terms in a different sense than we use them when talking about surface and volume charge densities, which have dimensions C/m$^2$ and C/m$^3$. For example, to obtain a current from a volume current density, we multiply it by an area.

---

**Figure 6.21.** At a sheet of surface current there must be a change in the parallel component of $B$ from one side to the other.
The magnetic field

thin, in any case, compared with the scale of any field variation.) The current enclosed by the rectangle is just $\mathcal{J}l$. Hence Eq. (6.19) yields

$$l(B_z^+ - B_z^-) = \mu_0 \mathcal{J}l,$$

or

$$B_z^+ - B_z^- = \mu_0 \mathcal{J}. \quad (6.58)$$

A current sheet of density $\mathcal{J}$ gives rise to a jump in the component of $\mathbf{B}$ that is parallel to the surface and perpendicular to $\mathcal{J}$. This may remind you of the change in electric field at a sheet of charge. There, the perpendicular component of $\mathbf{E}$ is discontinuous, the magnitude of the jump depending on the density of surface charge.

If the sheet is the only current source we have, then of course the field is symmetrical about the sheet; $B_z^+$ is $\mu_0 \mathcal{J}/2$, and $B_z^-$ is $-\mu_0 \mathcal{J}/2$. This is shown in Fig. 6.22(a). Some other situations, in which the effect of the current sheet is superposed on a field already present from another source, are shown in Fig. 6.22(b) and (c). Suppose there are two sheets carrying equal and opposite surface currents, as shown in cross section in Fig. 6.23, with no other sources around. The direction of current flow is perpendicular to the plane of the paper, out on the left and in on the right. The field between the sheets is $\mu_0 \mathcal{J}$, and there is no field at all outside. Something like this is found when current is carried by two parallel ribbons or slabs, close together compared with their width, as sketched in Fig. 6.24. Often bus bars for distributing heavy currents in power stations are of this form.

**Example (Field from a cylinder of current)** A cylindrical shell has radius $R$ and carries uniformly distributed current $I$ parallel to its axis. Find the magnetic field outside the shell, an infinitesimal distance away from it. Do this in the following way.

(a) Slice the shell into infinitely long “rods” parallel to the axis, and then integrate the field contributions from all the rods. You should obtain $B = \mu_0 I/2\pi R$. However, this is not the correct field, because we know from Ampère’s law that the field outside a wire (or a cylinder) takes the form of $\mu_0 I/2\pi r$. And $r = R$ here.

(b) What is wrong with the above reasoning? Explain how to modify it to obtain the correct result of $\mu_0 I/2\pi R$. Hint: You could have very well performed the above integral in an effort to obtain the magnetic field an infinitesimal distance inside the cylinder, where we know the field is zero.

**Solution**

(a) A cross section of the cylinder is shown in Fig. 6.25. A small piece of the circumference of the circle represents a rod pointing into and out of the page. Let the rods be parameterized by the angle $\theta$ relative to the point $P$ at which we are calculating the field. If a rod subtends an angle $d\theta$, then it contains a fraction $d\theta/2\pi$ of the total current $I$. So the current in the rod is $I(d\theta/2\pi)$. The rod is a distance $2R \sin(\theta/2)$ from $P$, which is infinitesimally close to the top of the cylinder.
6.6 Change in $B$ at a current sheet

If the current heads into the page, then the field due to the rod shown is directed up and to the right at $P$. Only the horizontal (tangential) component of this field survives, because the vertical component cancels with that from the corresponding rod on the left side of the cylinder. This brings in a factor of $\sin(\theta/2)$, as you can verify. Using the fact that the field from a straight rod takes the form of $\mu_0 I / 2\pi r$, we find that the field at point $P$ is directed to the right and has magnitude (apparently) equal to

$$B = 2\int_0^\pi \frac{\mu_0 I d\theta / 2\pi}{2\pi (2R \sin(\theta/2))} \sin(\theta/2) = \frac{\mu_0 I}{4\pi^2 R} \int_0^\pi d\theta = \frac{\mu_0 I}{4\pi R}. \quad (6.59)$$

(b) (You should try to solve this on your own before reading further. You may want to take a look at Exercise 1.67.) As noted in the statement of the problem, it is no surprise that the above result is incorrect, because the same calculation would supposedly yield the field just inside the cylinder too. But we know that the field there is zero. The calculation does, however, yield the next best thing, namely the average of the correct fields inside and outside (zero and $\mu_0 I / 2\pi R$). We’ll see why shortly.

The reason why the calculation is invalid is that it doesn’t correctly describe the field due to rods on the cylinder very close to the given point $P$, that is, for rods characterized by $\theta \approx 0$. It is incorrect for two reasons. The closeup view in Fig. 6.26 (with an exaggerated distance from $P$ to the cylinder, for clarity) shows that the distance from a rod to $P$ is not equal to $2R \sin(\theta/2)$. Additionally, it shows that the field at $P$ does not point perpendicular to the line from the rod to the top of the cylinder. It points more horizontally, so the extra factor of $\sin(\theta/2)$ in Eq. (6.59) is not correct.

What is true is that if we remove a thin strip from the top of the cylinder (so we now have a gap in the circle representing the cross-sectional view), then the above integral is valid for the remaining part of the cylinder. The strip contributes negligibly to the integral in Eq. (6.59) (assuming it subtends a very small angle), so we can say that the field due to the remaining part of the cylinder is equal to the above result of $\mu_0 I / 4\pi R$. By superposition, the total field due to the entire cylinder is this field of $\mu_0 I / 4\pi R$ plus the field due to the thin strip. But if the point in question is infinitesimally close to the cylinder, then the thin strip looks like an infinite sheet of current, the field of which we know is $\mu_0 J / 2 = \mu_0 (I / 2\pi R) / 2 = \mu_0 I / 4\pi R$.

The desired total field is then

$$B_{\text{outside}} = B_{\text{cylinder minus strip}} + B_{\text{strip}} = \frac{\mu_0 I}{4\pi R} + \frac{\mu_0 I}{2\pi R} = \frac{\mu_0 I}{2\pi R}. \quad (6.60)$$

The relative sign here is indeed a plus sign, because right above the strip, the strip’s field points to the right, which is the same as the direction of the field due to the rest of the cylinder. By superposition we also obtain the correct field just inside the cylinder:

$$B_{\text{inside}} = B_{\text{cylinder minus strip}} - B_{\text{strip}} = \frac{\mu_0 I}{4\pi R} - \frac{\mu_0 I}{2\pi R} = 0. \quad (6.61)$$

The relative minus sign comes from the fact that right below the strip, the strip’s field points to the left. For an alternative way of solving this problem, see Bose and Scott (1985).
The change in $\mathbf{B}$ from one side of a sheet to the other takes place within the sheet, as we already remarked in connection with Fig. 6.19. For the same $\mathbf{J}$, the thinner the sheet, the more abrupt the transition. We looked at a situation very much like this in Chapter 1 when we examined the discontinuity in the perpendicular component of $\mathbf{E}$ that occurs at a sheet of surface charge. It was instructive then to ask about the force on the surface charge, and we shall ask a similar question here.

Consider a square portion of the sheet, a length $\ell$ on a side (a rectangle would work fine too). The current included is equal to $J \ell$, the length of current path is $\ell$, and the average field that acts on this current is $(B_z^+ + B_z^-)/2$. The force on a length $\ell$ of current-carrying wire equals $IB\ell$ (see Eq. (6.14)), so the force on this portion of the current distribution is

$$\text{force on } \ell^2 \text{ of sheet} = IB_{\text{avg}}\ell = (J \ell) \left( \frac{B_z^+ + B_z^-}{2} \right) \ell. \quad (6.62)$$

In view of Eq. (6.58), we can substitute $(B_z^+ - B_z^-)/\mu_0$ for $J$, so that the force per unit area can be expressed in this way:

$$\text{force per unit area} = \left( \frac{B_z^+ - B_z^-}{\mu_0} \right) \left( \frac{B_z^+ + B_z^-}{2} \right) = \frac{1}{2\mu_0} \left[ (B_z^+)^2 - (B_z^-)^2 \right]. \quad (6.63)$$

The force is perpendicular to the surface and proportional to the area, like the stress caused by hydrostatic pressure. To make sure of the sign, we can figure out the direction of the force in a particular case, such as that in Fig. 6.23. The force is outward on each conductor. It is as if the high-field region were the region of high pressure. The repulsion of any two conductors carrying current in opposite directions, as in Fig. 6.24, can be seen as an example of that.

We have been considering an infinite flat sheet, but things are much the same in the immediate neighborhood of any surface where there is a change in $\mathbf{B}$. Wherever the component of $\mathbf{B}$ parallel to the surface changes from $B_1$ to $B_2$, from one side of the surface to the other, we may conclude not only that there is a sheet of current flowing in the surface, but also that the surface must be under a perpendicular stress of $(B_1^2 - B_2^2)/2\mu_0$, measured in N/m$^2$. This is one of the controlling principles in magnetohydrodynamics, the study of electrically conducting fluids, a subject of interest both to electrical engineers and to astrophysicists.

### 6.7 How the fields transform

A sheet of surface charge, if it is moving parallel to itself, constitutes a surface current. If we have a uniform charge density of $\sigma$ on the surface,
with the surface itself sliding along at speed $v$, the surface current density is just $J = \sigma v$. This is true because the area that slides past a transverse line with length $\ell$ during time $dt$ is $(v dt)\ell$, which yields a current of $\sigma (v dt)\ell / dt = (\sigma v)\ell$. And since the current is also $J\ell$ by definition, we have $J = \sigma v$. This simple idea of a sliding surface will help us to see how the electric and magnetic field quantities must change when we transform from one inertial frame of reference to another. We will deal first with the transverse fields, and then with the longitudinal fields.

Let’s imagine two plane sheets of surface charge, parallel to the $xz$ plane and moving with speed $v_0$ as in Fig. 6.27. Again, we show fragments of surfaces only in the sketch; the surfaces are really infinite in extent. In the inertial frame $F$ with coordinates $x$, $y$, and $z$, where the sheets move with speed $v_0$, the density of surface charge is $\sigma$ on one sheet and $-\sigma$ on the other. Here $\sigma$ means the amount of charge within unit area when area is measured by observers stationary in $F$. (It is not the density of charge in the rest frame of the charges themselves, which would be smaller by $1/\gamma_0$.) In the frame $F$ the uniform electric field $E$ points in the positive $y$ direction, and Gauss’s law assures us, as usual, that its strength is

$$E_y = \frac{\sigma}{\epsilon_0}. \quad (6.64)$$

In this frame $F$ the sheets are both moving in the positive $x$ direction with speed $v_0$, so that we have a pair of current sheets. The density of surface current is $J_x = \sigma v_0$ in one sheet, the negative of that in the other. As in the arrangement in Fig. 6.23, the field between two such current sheets is

$$B_z = \mu_0 J_x = \mu_0 \sigma v_0. \quad (6.65)$$

The inertial frame $F'$ is one that moves, as seen from $F$, with a speed $v$ in the positive $x$ direction. What fields will an observer in $F'$ measure? To answer this we need only find out what the sources look like in $F'$.

In $F'$ the $x'$ velocity of the charge-bearing sheets is $v'_0$, given by the velocity addition formula

$$v'_0 = \frac{v_0 - v}{1 - v_0 v/c^2} = \frac{\beta_0 - \beta}{1 - \beta_0 \beta}. \quad (6.66)$$

There is a different Lorentz contraction of the charge density in this frame, exactly as in our earlier example of the moving line charge in Section 5.9. We can repeat the argument we used then: the density in the rest frame of the charges themselves is $\sigma (1 - v_0^2/c^2)^{1/2}$, or $\sigma/\gamma_0$, and therefore the density of surface charge in the frame $F'$ is

$$\sigma' = \frac{\sigma}{\gamma_0} \frac{v'_0}{\gamma_0}. \quad (6.67)$$

Figure 6.27.
(a) As observed in frame $F$, the surface charge density is $\sigma$ and the surface current density is $\sigma v_0$. (b) Frame $F'$ moves in the $x$ direction with speed $v$ as seen from $F$. In $F'$ the surface charge density is $\sigma'$ and the current density is $\sigma' v'_0$. 

As usual, $\gamma'_0$ stands for $(1 - v_0^2/c^2)^{-1/2}$. Using Eq. (6.66), you can show that $\gamma'_0 = \gamma_0 \gamma(1 - \beta_0 \beta)$. Hence,

$$\sigma' = \sigma \gamma(1 - \beta_0 \beta).$$  \hfill (6.68)

The surface current density in the frame $F'$ is (charge density) $\times$ (charge velocity):

$$\mathcal{J}' = \sigma' v'_0 = \sigma \gamma(1 - \beta_0 \beta) \cdot c \frac{\beta_0 - \beta}{1 - \beta_0 \beta} = \sigma \gamma (v_0 - v).$$  \hfill (6.69)

We now know how the sources appear in frame $F'$, so we know what the fields in that frame must be. In saying this, we are again invoking the postulate of relativity. The laws of physics must be the same in all inertial frames, and that includes the formulas connecting electric field with surface charge density, and magnetic field with surface current density. It follows then that

$$E'_y = \frac{\sigma'}{\varepsilon_0} = \gamma \left[ \frac{\sigma}{\varepsilon_0} - \frac{\sigma}{\varepsilon_0} \frac{v_0}{c} \right] = \gamma \left[ \frac{\sigma}{\varepsilon_0} - \frac{v}{\mu_0 \varepsilon_0 c^2} \cdot \mu_0 \sigma v_0 \right],$$

$$B'_z = \mu_0 \mathcal{J}' = \gamma \left[ \mu_0 \sigma v_0 - \mu_0 \sigma v \right] = \gamma \left[ \mu_0 \sigma v_0 - \mu_0 \varepsilon_0 v \cdot \frac{\sigma}{\varepsilon_0} \right].$$  \hfill (6.70)

(These expressions might look a bit scary, but don’t worry, they’ll simplify!) We have chosen to write $E'_y$ and $B'_z$ in this way because if we look back at the values of $E_y$ and $B_z$ in Eqs. (6.64) and (6.65), we see that our result can be written as follows:

$$E'_y = \gamma \left( E_y - \frac{v}{\mu_0 \varepsilon_0 c^2} \cdot B_z \right),$$

$$B'_z = \gamma \left( B_z - \mu_0 \varepsilon_0 v \cdot E_y \right).$$  \hfill (6.71)

We can further simplify these expressions by using the relation $1/\mu_0 \varepsilon_0 = c^2$ from Eq. (6.8). We finally obtain

$$E'_y = \gamma \left( E_y - v B_z \right),$$

$$B'_z = \gamma \left( B_z - \frac{v}{c^2} E_y \right),$$  \hfill (6.72)

or equivalently

$$E'_y = \gamma \left( E_y - \beta (c B_z) \right),$$

$$c B'_z = \gamma \left( (c B_z) - \beta E_y \right).$$  \hfill (6.73)

You will note that these are exactly the same Lorentz transformations that apply to $x$ and $t$ (see Eq. (G.2) in Appendix G). They are symmetric in $E_y$ and $c B_z$. 

---

The magnetic field
If the sandwich of current sheets had been oriented parallel to the xy plane instead of the xz plane, we would have obtained relations connecting $E'_z$ with $E_z$ and $B_x$, and connecting $B'_y$ with $B_y$ and $E_z$. Of course, they would have the same form as the relations above, but if you trace the directions through, you will find that there are differences in sign, following from the rules for the direction of $B$.

Now we must learn how the field components in the direction of motion change. We discovered in Section 5.5 that a longitudinal component of $E$ has the same magnitude in the two frames. That this is true also of a longitudinal component of $B$ can be seen as follows. Suppose a longitudinal component of $B$, a $B_z$ component in the arrangement in Fig. 6.27, is produced by a solenoid around the x axis in frame $F$ (at rest in $F$). The field strength inside a solenoid, as we know from Eq. (6.57), depends only on the current in the wire, $I$, which is charge per second, and $n$, the number of turns of wire per meter of axial length. In the frame $F'$ the solenoid will be Lorentz-contracted, so the number of turns per meter in that frame will be greater. But the current, as reckoned by observers in $F'$, will be reduced, since, from their point of view, the $F$ observers who measured the current by counting the number of electrons passing a point on the wire, per second, were using a slow-running watch. The time dilation just cancels the length contraction in the product $nI$. Indeed any quantity of the dimensions (longitudinal length)$^{-1} \times$ (time)$^{-1}$ is unchanged in a Lorentz transformation. So $B'_x = B_x$.

Remember the point made early in Chapter 5, in the discussion following Eq. (5.6): the transformation properties of the field are local properties. The values of $E$ and $B$ at some space-time point in one frame must uniquely determine the field components observed in any other frame at that same space-time point. Therefore the fact that we have used an especially simple kind of source (the parallel uniformly charged sheets, or the solenoid) in our derivation in no way compromises the generality of our result. We have in fact arrived at the general laws for the transformation of all components of the electric and magnetic field, of whatever origin or configuration.

We give below the full list of transformations. All primed quantities are measured in the frame $F'$, which is moving in the positive $x$ direction with speed $v$ as seen from $F$. Unprimed quantities are the numbers that are the results of measurement in $F$. As usual, $\beta$ stands for $v/c$ and $\gamma$ for $(1 - \beta^2)^{-1/2}$.

\[
\begin{align*}
E'_x &= E_x \quad E'_y = \gamma (E_y - vB_z) \quad E'_z = \gamma (E_z + vB_y) \\
B'_x &= B_x \quad B'_y = \gamma (B_y + (v/c^2)E_z) \quad B'_z = \gamma (B_z - (v/c^2)E_y)
\end{align*}
\]

(6.74)

When these equations are written in the alternative form given in Eq. (6.73), the symmetry between $E$ and $cB$ is evident. If the printer
The magnetic field

had mistakenly interchanged $E$’s with $cB$’s, and $y$’s with $z$’s, the equations would come out exactly the same. Certainly magnetic phenomena as we find them in Nature are distinctly different from electrical phenomena. The world around us is by no means symmetrical with respect to electricity and magnetism. Nevertheless, with the sources out of the picture, we find that the fields themselves, $E$ and $cB$, are connected to one another in a highly symmetrical way.

It appears too that the electric and magnetic fields are in some sense aspects, or components, of a single entity. We can speak of the electromagnetic field, and we may think of $E_x$, $E_y$, $E_z$, $cB_x$, $cB_y$, and $cB_z$ as six components of the electromagnetic field. The same field viewed in different inertial frames will be represented by different sets of values for these components, somewhat as a vector is represented by different components in different coordinate systems rotated with respect to one another. However, the electromagnetic field so conceived is not a vector, mathematically speaking, but rather something called a tensor. The totality of the equations in the box on page 309 forms the prescription for transforming the components of such a tensor when we shift from one inertial frame to another. We are not going to develop that mathematical language here. In fact, we shall return now to our old way of talking about the electric field as a vector field, and the magnetic field as another vector field coupled to the first in a manner to be explored further in Chapter 7. To follow up on this brief hint of the unity of the electromagnetic field as represented in four-dimensional space-time, you will have to wait for a more advanced course.

We can express the transformation of the fields, Eq. (6.74), in a more elegant way which is often useful. Let $v$ be the velocity of a frame $F'$ as seen from a frame $F$. We can always resolve the fields in both $F$ and $F'$ into vectors parallel to and perpendicular to, respectively, the direction of $v$. Thus, using an obvious notation:

$$E = E \parallel + E \perp, \quad E' = E' \parallel + E' \perp,$$
$$B = B \parallel + B \perp, \quad B' = B' \parallel + B' \perp. \quad (6.75)$$

Then the transformation can be written like this (as you can verify):

$$E' \parallel = E \parallel, \quad E' \perp = \gamma (E \perp + v \times B \perp),$$
$$B' \parallel = B \parallel, \quad B' \perp = \gamma (B \perp - (v/c^2) \times E \perp) \quad (6.76)$$

($v$ is the velocity of $F'$ with respect to $F$)

In the special case that led us to Eq. (6.72), $v$ was $v\hat{x}$, $E \perp$ was $(\sigma/\epsilon_0)\hat{y}$, and $B_\perp$ was $\mu_0\sigma_0\hat{z}$. You can check that these vectors turn the “$\perp$” equations in Eq. (6.76) into the equations in Eq. (6.70); you will need to use $1/\mu_0\epsilon_0 = c^2$. 
In Gaussian units, with $E$ in statvolts/cm and $B$ in gauss, the Lorentz transformation of the fields reads as follows (with $\beta \equiv v/c$):

$$
\begin{align*}
E'_\parallel &= E_\parallel \\
E'_\perp &= \gamma \left( E_\perp + \beta \times B_\perp \right) \\
B'_\parallel &= B_\parallel \\
B'_\perp &= \gamma \left( B_\perp - \beta \times E_\perp \right)
\end{align*}
$$

(6.77)

You can derive these relations by working through the above procedure in Gaussian units. But a quicker method is to note that, when changing formulas from SI to Gaussian units, we must replace $\epsilon_0$ with $1/4\pi$ (by looking at the analogous expressions for Coulomb’s law, or equivalently Gauss’s law), and $\mu_0$ with $4\pi/c$ (by looking at the analogous expressions for Ampère’s law). The product $\mu_0 \epsilon_0$ therefore gets replaced with $1/c$, and Eq. (6.71) then leads to Eq. (6.77).

One advantage of the Gaussian system of units is that the transformations in Eq. (6.77) are more symmetrical than those in Eq. (6.76). This can be traced to the fact that $E$ and $B$ have the same units in the Gaussian system. In the SI system, unfortunately, the use of different units for $E$ and $B$ (due to the definition of $B$ in Eq. (6.1)) tends to obscure the essential electromagnetic symmetry of the vacuum. The electric and magnetic fields are after all components of one tensor. The Lorentz transformation is something like a rotation, turning $E$ partly into $B'$, and $B$ partly into $E'$. It seems quite natural and appropriate that the only parameter in Eq. (6.77) is the dimensionless ratio $\beta$. To draw an analogy that is not altogether unfair, imagine that it has been decreed that east–west displacement components must be expressed in meters while north–south components are to be in feet. The transformation effecting a rotation of coordinate axes would be, to say the least, aesthetically unappealing. Nor is symmetry restored to Eq. (6.76) when $B$ is replaced, as is often done, by a vector $\mathbf{H}$, which we shall meet in Chapter 11, and which in the vacuum is simply $B/\mu_0$.

Having said this, however, we should note that there isn’t anything disastrous about the extra factor of $c$ that appears in the SI Lorentz transformation. The Lorentz symmetry between $E$ and $cB$ is similar to the Lorentz symmetry between $x$ and $ct$. The coordinates $x$ and $t$ have different dimensions, but are still related by a Lorentz transformation with an extra factor of $c$ thrown in.

---

**Example (Stationary charge and rod)** A charge $q$ is at rest a distance $r$ from a long rod with linear charge density $\lambda$, as shown in Fig. 6.28. The charges in the rod are also at rest. The electric field due to the rod takes the standard form of $E = \lambda/2\pi \epsilon_0 r$, so the force on the charge $q$ in the lab frame is simply $F = qE = q\lambda/2\pi \epsilon_0 r$. This force is repulsive, assuming $q$ and $\lambda$ have the same sign.

---

**Figure 6.28.** A point charge at rest with respect to a charged rod.
Now consider the setup in the frame that moves to the left with speed $v$. In this frame both the charge $q$ and the charges in the rod move to the right with speed $v$. What is the force on the charge $q$ in this new frame? Solve this in three different ways.

(a) Transform the force from the lab frame to the new frame, without caring about where it comes from in the new frame.

(b) Directly calculate the electric and magnetic forces in the new frame, by considering the charges in the rod.

(c) Transform the fields using the Lorentz transformations.

Solution

(a) The force on a particle is always largest in the rest frame of the particle. It is smaller in any other frame by the $\gamma$ factor associated with the speed $v$ of the particle. The force in the particle frame (the lab frame) is $q\lambda/2\pi\epsilon_0 r$, so the force in the new frame is $q\lambda/2\gamma\pi\epsilon_0 r$.

(b) In the new frame (call it $F'$), the linear charge density in the rod is increased to $\gamma\lambda$, due to length contraction. So the electric field is $E' = \gamma\lambda/(2\pi\epsilon_0 r)$. This field produces a repulsive electric force of $F_E = \gamma q\lambda/2\pi\epsilon_0 r$.

In $F'$ the current produced by the rod is the density times the speed, so $I = (\gamma\lambda)v$. The magnetic field is then $B' = \mu_0 I/2\pi r = \mu_0\gamma\lambda v/2\pi r$, directed into the page in Fig. 6.29 (assuming $\lambda$ is positive). The magnetic force is therefore attractive and has magnitude (using $\mu_0 = 1/\epsilon_0 c^2$)

$$F_B = qvB' = qv \cdot \mu_0 \gamma\lambda v/(2\pi r) = \gamma q\lambda v^2/2\pi\epsilon_0 r c^2.$$ (6.78)

The net repulsive force acting on the charge $q$ in the new frame is therefore

$$F_E - F_B = \gamma q\lambda/2\pi\epsilon_0 r - \gamma q\lambda v^2/2\pi\epsilon_0 r c^2 = \gamma q\lambda/2\pi\epsilon_0 r \left(1 - v^2/c^2\right) = q\lambda/2\gamma\pi\epsilon_0 r,$$ (6.79)

where we have used $1 - v^2/c^2 \equiv 1/\gamma^2$. This net force agrees with the result in part (a).

(c) In the lab frame, the charges in the rod aren’t moving, so $E_{\perp}$ is the only nonzero field in the Lorentz transformations in Eq. (6.76). It is directed away from the rod with magnitude $\lambda/(2\pi\epsilon_0 r)$. Equation (6.76) immediately gives the electric field in the new frame as $E'_{\perp} = \gamma E_{\perp}$. So $E'_{\perp}$ has magnitude $E'_{\perp} = \gamma\lambda/(2\pi\epsilon_0 r)$ and is directed away from the rod, in agreement with the electric field we found in part (b).

Equation (6.76) gives the magnetic field in the new frame as $B'_{\perp} = -\gamma(v/c^2) \times E_{\perp}$. The velocity $v$ of $F'$ with respect to the lab frame $F$ points to the left with magnitude $v$. We therefore find that $B'_{\perp}$ points into the page with magnitude $B'_{\perp} = \gamma(v/c^2)(\lambda/2\pi\epsilon_0 r)$. In terms of $\mu_0 = 1/\epsilon_0 c^2$, this can be written as $B'_{\perp} = \mu_0\gamma\lambda v/2\pi r$, in agreement with the magnetic field we found in part (b). We therefore arrive at the same net force, $F_E - F_B$, as in part (b).
There is a remarkably simple relation between the electric and magnetic field vectors in a special but important class of cases. Suppose a frame exists – let’s call it the unprimed frame – in which \( \mathbf{B} \) is zero in some region (as in the above example). Then in any other frame \( F' \) that moves with velocity \( \mathbf{v} \) relative to that special frame, we have, according to Eq. (6.76),

\[
\begin{align*}
\mathbf{E}'_\parallel &= \mathbf{E}_\parallel, \\
\mathbf{E}'_\perp &= \gamma \mathbf{E}_\perp, \\
\mathbf{B}'_\parallel &= 0, \\
\mathbf{B}'_\perp &= -\gamma \left( \frac{\mathbf{v}}{c^2} \right) \times \mathbf{E}_\perp.
\end{align*}
\]

(6.80)

In the last of these equations we can replace \( \mathbf{B}'_\perp \) with \( \mathbf{B}' \), because \( \mathbf{B}'_\parallel = 0 \). We can also replace \( \gamma \mathbf{E}_\perp \) with \( \mathbf{E}'_\perp \), which we can in turn replace with \( \mathbf{E}' \) because \( \mathbf{v} \times \mathbf{E}'_\parallel = 0 \) (since \( \mathbf{E}'_\parallel \) is parallel to \( \mathbf{v} \) by definition). The last equation therefore becomes a simple relation between the full \( \mathbf{E}' \) and \( \mathbf{B}' \) fields:

\[
\mathbf{B}' = -\left( \frac{\mathbf{v}}{c^2} \right) \times \mathbf{E}' \quad \text{(if } \mathbf{B} = 0 \text{ in one frame).}
\]

(6.81)

This holds in every frame if \( \mathbf{B} = 0 \) in one frame. Remember that \( \mathbf{v} \) is the velocity of the frame in question (the primed frame) with respect to the special frame in which \( \mathbf{B} = 0 \).

In the same way, we can deduce from Eq. (6.76) that, if there exists a frame in which \( \mathbf{E} = 0 \), then in any other frame

\[
\mathbf{E}' = \mathbf{v} \times \mathbf{B}' \quad \text{(if } \mathbf{E} = 0 \text{ in one frame).}
\]

(6.82)

As before, \( \mathbf{v} \) is the velocity of the frame \( F' \) with respect to the special frame \( F \) in which, in this case, \( \mathbf{E} = 0 \).

Because Eqs. (6.81) and (6.82) involve only quantities measured in the same frame of reference, they are easy to apply, whenever the restriction is met, to fields that vary in space. A good example is the field of a point charge \( q \) moving with constant velocity, the problem studied in Chapter 5. Take the unprimed frame to be the frame in which the charge is at rest. In this frame, of course, there is no magnetic field. Equation (6.81) tells us that in the lab frame, where we find the charge moving with speed \( \mathbf{v} \), there must be a magnetic field perpendicular to the electric field and to the direction of motion. We have already worked out the exact form of the electric field in this frame: we know the field is radial from the instantaneous position of the charge, with a magnitude given by Eq. (5.15). The magnetic field lines must be circles around the direction of motion, as indicated crudely in Fig. 6.30. When the velocity of the charge is high (\( \mathbf{v} \approx c \)), so that \( \gamma \gg 1 \), the radial “spokes” that are the electric field lines are folded together into a thin disk. The circular magnetic field lines are likewise concentrated in this disk. The magnitude of \( \mathbf{B} \) is then nearly equal to the magnitude of \( \mathbf{E}/c \). That is, the magnitude of the magnetic field in tesla is almost exactly the same as \( 1/c \) times the magnitude of the electric field in volts/meter, at the same point and instant of time.
We have come a long way from Coulomb’s law in the last two chapters. Yet with each step we have only been following out consistently the requirements of relativity and of the invariance of electric charge. We can begin to see that the existence of the magnetic field and its curiously symmetrical relationship to the electric field is a necessary consequence of these general principles. We remind the reader again that this was not at all the historical order of discovery and elucidation of the laws of electromagnetism. One aspect of the coupling between the electric and magnetic fields, which is implicit in Eq. (6.74), came to light in Michael Faraday’s experiments with changing electric currents, which will be described in Chapter 7. That was 75 years before Einstein, in his epochal paper of 1905, first wrote out our Eq. (6.74).

6.8 Rowland’s experiment
As we remarked in Section 5.9, it was not obvious 150 years ago that a current flowing in a wire and a moving electrically charged object are essentially alike as sources of magnetic field. The unified view of electricity and magnetism that was then emerging from Maxwell’s work suggested that any moving charge ought to cause a magnetic field, but experimental proof was hard to come by.

That the motion of an electrostatically charged sheet produces a magnetic field was first demonstrated by Henry Rowland, the great American physicist renowned for his perfection of the diffraction grating. Rowland made many ingenious and accurate electrical measurements, but none that taxed his experimental virtuosity as severely as the detection and measurement of the magnetic field of a rotating charged disk. The field to be detected was something like $10^{-5}$ of the earth’s field in magnitude – a formidable experiment, even with today’s instruments! In Fig. 6.31, you will see a sketch of Rowland’s apparatus and a reproduction of the first page of the paper in which he described his experiment. Ten years before Hertz’s discovery of electromagnetic waves, Rowland’s result gave independent, if less dramatic, support to Maxwell’s theory of the electromagnetic field.

6.9 Electrical conduction in a magnetic field: the Hall effect
When a current flows in a conductor in the presence of a magnetic field, the force $qv \times B$ acts directly on the moving charge carriers. Yet we observe a force on the conductor as a whole. Let’s see how this comes about. Figure 6.32(a) shows a section of a metal bar in which a steady current is flowing. Driven by a field $E$, electrons are drifting to the left with average speed $\bar{v}$, which has the same meaning as the $\bar{u}$ in our discussion of conduction in Chapter 4. The conduction electrons are indicated, very schematically, by the gray dots. The black dots are the positive
ON THE MAGNETIC EFFECT OF ELECTRIC CONVECTION

[American Journal of Science [3], XV, 30–38, 1878]

The experiments described in this paper were made with a view of determining whether or not an electrified body in motion produces magnetic effects. There seems to be no theoretical ground upon which we can settle the question, seeing that the magnetic action of a conducted electric current may be ascribed to some mutual action between the conductor and the current. Hence an experiment is of value. Professor Maxwell, in his ‘Treatise on Electricity,’ Art. 770, has computed the magnetic action of a moving electrified surface, but that the action exists has not yet been proved experimentally or theoretically.

The apparatus employed consisted of a vulcanite disc 21·1 centimetres in diameter and .5 centimetre thick which could be made to revolve around a vertical axis with a velocity of 61· turns per second. On either side of the disc at a distance of .6 cm. were fixed glass plates having a diameter of 38·9 cm. and a hole in the centre of 7·8 cm. The vulcanite disc was gilded on both sides and the glass plates had an annular ring of gilt on one side, the outside and inside diameters being 24·0 cm. and 8·9 cm. respectively. The gilt sides could be turned toward or from the revolving disc but were usually turned toward it so that the problem might be calculated more readily and there should be no uncertainty as to the electrification. The outside plates were usually connected with the earth; and the inside disc with an electric battery, by means of a point which approached within one-third of a millimetre of the edge and turned toward it. As the edge was broad, the point would not discharge unless there was a difference of potential between it and the edge. Between the electric battery and the disc, . . .

1 The experiments described were made in the laboratory of the Berlin University through the kindness of Professor Helmholtz, to whose advice they are greatly indebted for their completeness. The idea of the experiment first occurred to me in 1868 and was recorded in a note book of that date.
Figure 6.32.
(a) A current flows in a metal bar. Only a short section of the bar is shown. Conduction electrons are indicated (not in true size and number!) by gray dots, positive ions of the crystal lattice by black dots. The arrows indicate the average velocity \( \mathbf{v} \) of the electrons.
(b) A magnetic field is applied in the \( x \) direction, causing (at first) a downward deflection of the moving electrons. (c) The altered charge distribution makes a transverse electric field \( E_t \). In this field the stationary positive ions experience a downward force.

The magnetic field

ions which form the rigid framework of the solid metal bar. Since the electrons are negative, we have a current in the positive \( y \) direction. The current density \( \mathbf{J} \) and the field \( \mathbf{E} \) are related by the conductivity of the metal, \( \sigma \), as usual: \( \mathbf{J} = \sigma \mathbf{E} \). There is no magnetic field in Fig. 6.32(a) except that of the current itself, which we shall ignore.

Now an external field \( \mathbf{B} \) in the \( x \) direction is switched on. The state of motion immediately thereafter is shown in Fig. 6.32(b). The electrons are being deflected downward. But since they cannot escape at the bottom of the bar, they simply pile up there, until the surplus of negative charge at the bottom of the bar and the corresponding excess of positive charge at the top create a downward transverse electric field \( E_t \) in which the upward force, of magnitude \( eE_t \), exactly balances the downward force \( e\mathbf{v}\mathbf{B} \). In the steady state (which is attained very quickly!) the average motion is horizontal again, and there exists in the interior of the metal this transverse electric field \( E_t \), as observed in coordinates fixed in the metal lattice (Fig. 6.32(c)). This field causes a downward force on the positive ions. That is how the force, \( -e\mathbf{v} \times \mathbf{B} \), on the electrons is passed on to the solid bar. The bar, of course, pushes on whatever is holding it.

The condition for zero average transverse force on the moving charge carriers is

\[
E_t + \mathbf{v} \times \mathbf{B} = 0. \quad (6.83)
\]

Suppose there are \( n \) mobile charge carriers per \( m^3 \) and, to be more general, denote the charge of each by \( q \). Then the current density \( \mathbf{J} \) is \( nq\mathbf{v} \). If we now substitute \( \mathbf{J}/nq \) for \( \mathbf{v} \) in Eq. (6.83), we can relate the transverse field \( E_t \) to the directly measurable quantities \( \mathbf{J} \) and \( \mathbf{B} \):

\[
E_t = -\frac{\mathbf{J} \times \mathbf{B}}{nq}. \quad (6.84)
\]

For electrons \( q = -e \), so \( E_t \) has in that case the direction of \( \mathbf{J} \times \mathbf{B} \), as it does in Fig. 6.32(c).

The existence of the transverse field can easily be demonstrated. Wires are connected to points \( P_1 \) and \( P_2 \) on opposite edges of the bar (Fig. 6.33), the junction points being carefully located so that they are
at the same potential when current is flowing in the bar and \( B \) is zero. The wires are connected to a voltmeter. After the field \( B \) is turned on, \( P_1 \) and \( P_2 \) are no longer at the same potential. The potential difference is \( E_t \) times the width of the bar, and in the case illustrated \( P_1 \) is positive relative to \( P_2 \). A steady current will flow around the external circuit from \( P_1 \) to \( P_2 \), its magnitude determined by the resistance of the voltmeter. Note that the potential difference would be reversed if the current \( J \) consisted of positive carriers moving to the right rather than electrons moving to the left. This is true because the positive charge carriers would be deflected downward, just as the electrons were (because two things in the \( qv \times B \) force have switched signs, namely the \( q \) and the \( v \)). The field \( E_t \) would therefore have the opposite sign, being now directed upward.

Here for the first time we have an experiment that promises to tell us the sign of the charge carriers in a conductor.

The effect was discovered in 1879 by E. H. Hall, who was studying under Rowland at Johns Hopkins. In those days no one understood the mechanism of conduction in metals. The electron itself was unknown. It was hard to make much sense of the results. Generally the sign of the “Hall voltage” was consistent with conduction by negative carriers, but there were exceptions even to that. A complete understanding of the Hall effect in metallic conductors came only with the quantum theory of metals, about 50 years after Hall’s discovery.

The Hall effect has proved to be especially useful in the study of semiconductors. There it fulfills its promise to reveal directly both the concentration and the sign of the charge carriers. The \( n \)-type and \( p \)-type semiconductors described in Chapter 4 give Hall voltages of opposite sign, as we should expect. As the Hall voltage is proportional to \( B \), an appropriate semiconductor in the arrangement of Fig. 6.33 can serve, once calibrated, as a simple and compact device for measuring an unknown magnetic field. An example is described in Exercise 6.73.

**6.10 Applications**

A mass spectrometer is used to determine the chemical makeup of a substance. Its operation is based on the fact that if a particle moves perpendicular to a magnetic field, the radius of curvature of the circular path depends on the particle’s mass (see Exercise 6.29). In the spectrometer, molecules in a sample are first positively ionized, perhaps by bombarding them with electrons, which knocks electrons free. The ions are accelerated through a voltage difference and then sent through a magnetic field. Lighter ions have a smaller radius of curvature. (Even though lighter ions are accelerated to higher speeds, you can show that the magnetic field still bends them more compared with heavier ions.) By observing the final positions, the masses (or technically, the mass-to-charge ratios) of the various ions can be determined. Uses of mass spectroscopy include
forensics, drug testing, testing for contaminants in food, and determining the composition of the atmosphere of planets.

The image in an old television set (made prior to the early 2000s) is created by a cathode ray tube. The wide end of the tube is the television screen. Electrons are fired toward the screen and are deflected by the magnetic field produced by current-carrying coils. This current is varied in such a way that the impact point of the electrons on the screen traces out the entire screen (many times per second) via a sequence of horizontal lines. When the electrons hit the screen, they cause phosphor in the screen to emit light. A particular image appears, depending on which locations are illuminated; the intensity of the electron beam is modulated to create the desired shading at each point (black, white, or something in between). Color TVs have three different electron beams for the three primary colors.

Consider two coaxial solenoids, with their currents oriented the same way, separated by some distance in the longitudinal direction. The magnetic field lines will diverge as they leave one solenoid, then reach a maximum width halfway between the solenoids, and then converge as they approach the other solenoid. Consider a charged particle moving approximately in a circle perpendicular to the field lines (under the influence of the Lorentz force), while also drifting in the direction of the field lines. It turns out that the drifting motion will be reversed in regions where the field lines converge, provided that they converge quickly enough. The particle can therefore be trapped in what is called a magnetic bottle, bouncing back and forth between the ends. A bottle-type effect can be used, for example, to contain plasma in fusion experiments.

A magnetic bottle can also be created by the magnetic field due to a current ring. The ring effectively produces a curved magnetic bottle; the field lines expand near the plane of the ring and converge near the axis. This is basically what the magnetic field of the earth looks like. The regions of trapped charged particles are called the Van Allen belts. If particles approach the ends of the belt, that is, if they approach the earth’s atmosphere, the collisions with the air molecules cause the molecules to emit light. We know this light as the northern (or southern) lights, or alternatively as the aurora borealis (or australis). The various colors come from the different atomic transitions in oxygen and nitrogen. The sources of the charged particles are solar wind and cosmic rays. However, the source can also be man-made: a 1962 high-altitude hydrogen bomb test, code-named “Starfish Prime,” gave a wide area of the Pacific Ocean quite a light show. But at least people were “warned” about the test; the headline of the Honolulu Advertiser read, “N-Blast Tonight May Be Dazzling: Good View Likely.”

The earth’s atmosphere protects us significantly from solar wind (consisting mostly of protons and electrons) and from the steady background of cosmic rays (consisting mostly of protons). But the earth’s magnetic field also helps out. The charged particles are deflected away
from the earth by the Lorentz force from the magnetic field. In the event of a severe solar flare, however, a larger number of particles make it through the field, disrupting satellites and other electronics. Providing a way to shield astronauts from radiation is one of the main obstacles to extended space travel. Generating a magnetic field via currents in superconductors (see Appendix I) is a potential solution. Although this would certainly be expensive, propelling thick heavy shielding (the analog of the earth’s atmosphere) into space would also be very costly.

A railgun is a device that uses the $qv \times B$ Lorentz force, instead of an explosive, to accelerate a projectile. The gun consists of two parallel conducting rails with a conducting object spanning the gap between them. This object (which is the projectile, or perhaps a larger object holding the projectile) is free to slide along the rails. A power source sends current down one rail, across the projectile, and back along the other rail. The current in the rails produces a magnetic field, and you can use the right-hand rule to show that the resulting Lorentz force on the projectile is directed away from the power source. Large-scale rail guns can achieve projectile speeds of a few kilometers per second and have a range of hundreds of kilometers.

In one form of electric motor, a DC current flows through a coil that is free to rotate between the poles of a fixed magnet. The Lorentz force on the charges moving in the coil produces a torque (see Exercise 6.34), so the coil is made to rotate. (Alternatively, the coil behaves just like a magnet with north and south poles, and this magnet interacts with the field of the fixed magnet.) The coil can then apply a torque to whatever object the motor is attached to. There are both AC and DC motors, and different kinds of each. In a brushed DC motor, a commutator causes the current to change direction every half cycle, so that the torque is always in the same direction. If for some reason the motor gets stuck and stops rotating, the back emf (which we will learn about in Chapter 7) drops to zero. The current through the coil then increases, causing it to heat up. The accompanying smell will let you know that you are in the process of burning out your motor.

Solenoids containing superconducting wires can be used to create very large fields, on the order of 20 T. The superconducting wires allow a large current to flow with no resistive heating. A magnetic resonance imaging (MRI) machine uses the physics described in Appendix J to make images of the inside of your body. Its magnetic field is usually about one or two tesla. The magnets in the Large Hadron Collider at CERN are also superconducting. If a defect causes the circuit to become nonsuperconducting (a result known as a quench), then the circuit will heat up rapidly and a chain reaction of very bad things is likely. As we will see in Chapter 7, solenoids store energy, and this energy needs to go somewhere. In 2008 at CERN, a quench caused major damage and took the accelerator offline for a year.
The size of the magnetic field in a superconducting solenoid is limited by the fact that a superconducting wire can’t support magnetic fields or electric currents above certain critical values. The largest sustained magnetic fields in a laboratory are actually created with resistive conductors. The optimal design has the coils of the solenoid replaced by Bitter plates arranged in a helical pattern. These plates operate in basically the same manner as coils, but they allow for water cooling via well-placed holes. A Bitter magnet, which can achieve a field of about 35 T, requires a serious amount of water cooling – around 100 gallons per second!

If a coil of wire is wrapped around an iron core, the coil’s magnetic field is magnified by the iron, for reasons we will see in Chapter 11. The magnification factor can be 100 or 1000, or even larger. This magnification effect is used in relays, circuit breakers, junkyard magnets (all discussed below), and many other devices. The combination of a coil and an iron core (or even a coil without a core) is called an electromagnet. The main advantage of an electromagnet over a permanent magnet is that the field can be turned on and off.

An electric relay is a device that uses a small signal to switch on (or off) a larger signal. (So a relay and a transistor act in the same manner.) A small current in one circuit passes through an electromagnet. The magnetic field of this electromagnet pulls on a spring-mounted iron lever that closes (or opens) a second circuit. The power source in this second circuit is generally much larger than in the first. A relay is used, for example, in conjunction with a thermostat. A small current in the thermostat’s temperature sensor switches on a much larger current in the actual heating system, which involves, say, a hot-water pump. In years past, relays were used as telegraph repeaters. The relay took a weak incoming signal at the end of a long wire and automatically turned it into a strong outgoing signal, thereby eliminating the need for a human being to receive and retransmit the information.

A circuit breaker is similar to a fuse (see Section 4.12), in that it prevents the current in a circuit from becoming too large; 15 A or 20 A are typical thresholds for a household circuit breaker. If you are running many appliances in your home at the same time, the total current may be large enough to cause a wire somewhere inside a wall to overheat and start a fire. Similarly, a sustained short circuit will almost certainly cause a fire. One type of circuit breaker contains an electromagnet whose magnetic field pulls on an iron lever. The lever is held in place by a spring, but if the current in the electromagnet becomes sufficiently large, the force on the lever is large enough to pull it away from its resting position. This movement breaks the circuit in one way or another, depending on the design. The circuit breaker can be reset by simply flipping a switch that manually moves the lever back to its resting position. This should be contrasted with a fuse, which must be replaced each time it burns out.
**Junkyard magnets** can have fields in the 1 tesla range. This is fairly large, but the real key to the magnet’s strength is its large area, which is on the order of a square meter. If you’ve ever played with a small rare-earth magnet (with a field of around 1 tesla) and felt how strongly it can stick to things, imagine playing with one that is a meter in diameter. It’s no wonder it can pick up a car!

**Doorbells** (or at least those of the “ding dong” type) consist of a piston located inside a solenoid. Part of the piston is a permanent magnet. A spring holds the piston off to one side (say, to the left), where its left end rests on a sound bar. When the doorbell button is pressed, a circuit is completed and current flows through the solenoid. The resulting magnetic field pulls the piston through the solenoid and causes the right end to hit another sound bar located off to the right (this is the “ding” bar). As long as the button is held down, the piston stays there. But when the button is released, the current stops flowing in the solenoid, and the spring pushes the piston back to its initial position, where it strikes the left sound bar (the “dong” bar).

The **speakers** in your sound system are simple devices in principle, although they require a great deal of engineering to produce a quality sound. A speaker converts an electrical signal into sound waves. A coil of wire is located behind a movable cone (the main part of the speaker that you see, also called the diaphragm) and attached to its middle. The coil surrounds one pole of a permanent magnet, and is in turn surrounded by the other pole (imagine a stamp/cutter for making doughnuts; the coil is free to slide along the inner cylinder of the stamp). Depending on the direction of the current in the coil, the permanent magnet pushes the coil (and hence the cone) one way or the other. The movement of the cone produces the sound waves that travel to your ears. If the correct current (with the proper time-varying amplitude to control the volume, and frequency to control the pitch) is fed through the coil, the cone will oscillate in exactly the manner needed to produce the desired sound. Audible frequencies range from roughly 20 Hz to 20 kHz, so the oscillations will be quick. The necessary current originates in a **microphone**, which operates in the same way as a speaker, but in reverse. That is, a microphone converts sound waves into an electrical signal. There are many different types of microphones; we talked about one type in Section 3.9, and we will talk about another in Section 7.11, after we have covered electromagnetic induction.

**Maglev trains** (short for “magnetic levitation”) are vertically supported, laterally stabilized, and longitudinally accelerated by magnetic fields. The train has no contact with the track, which means that it can go faster than a conventional train; speeds can reach 500 km/hr, or 300 mph. There is also less wear and tear. There are two main types of maglev trains: electromagnetic suspension (EMS) and electrodynamic suspension (EDS). The EMS system makes use of both permanent magnets and electromagnets. The lateral motion of the train is unstable, so...
sensitive computerized correction is required. The EDS system makes use of electromagnets (in some cases involving superconductors), along with magnetic induction (discussed in Chapter 7). An advantage of EDS is that the lateral motion is stable, but a disadvantage is that the stray magnetic fields inside the cabin can be fairly large. Additionally, an EDS train requires a minimum speed to levitate, so wheels are needed at low speeds. The propulsion mechanism in both systems involves using alternating current to create magnetic fields that continually accelerate (or decelerate) the train. All maglev trains require a specially built track, which is a large impediment to adoption.

CHAPTER SUMMARY

- The Lorentz force on a charged particle in an electromagnetic field is
  \[ \mathbf{F} = q \mathbf{E} + q \mathbf{v} \times \mathbf{B}. \]  
  (6.85)

- The magnetic field due to a current in a long straight wire points in the tangential direction and has magnitude
  \[ B = \frac{I}{2\pi \epsilon_0 r c^2} = \frac{\mu_0 I}{2\pi r}, \]  
  (6.86)
  where
  \[ \mu_0 \equiv 4\pi \cdot 10^{-7} \frac{\text{kg m}}{\text{C}^2} \quad \text{and} \quad c^2 = \frac{1}{\mu_0 \epsilon_0}. \]  
  (6.87)
  If a wire carrying current \( I_2 \) lies perpendicular to a magnetic field \( B_1 \), then the magnitude of the force on a length \( l \) of the wire is \( F = I_2 B_1 l \). The SI and Gaussian units of magnetic field are the tesla and gauss, respectively. One tesla equals exactly \( 10^4 \) gauss.

- Ampère’s law in integral and differential form is
  \[ \int \mathbf{B} \cdot d\mathbf{s} = \mu_0 I \quad \iff \quad \text{curl} \mathbf{B} = \mu_0 \mathbf{J}. \]  
  (6.88)
  The magnetic field also satisfies
  \[ \text{div} \mathbf{B} = 0. \]  
  (6.89)
  This is the statement that there are no magnetic monopoles, or equivalently that magnetic field lines have no endings.

- The vector potential \( \mathbf{A} \) is defined by
  \[ \mathbf{B} = \text{curl} \mathbf{A}, \]  
  (6.90)
  which leads to \( \text{div} \mathbf{B} = 0 \) being identically true. Given the current density \( \mathbf{J} \), the vector potential can be found via
  \[ \mathbf{A} = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J} \, dv}{r} \quad \text{or} \quad \mathbf{A} = \frac{\mu_0 I}{4\pi} \int \frac{dl}{r} \]  
  (for a thin wire).
  (6.91)
The contribution to the magnetic field from a piece of a wire carrying current $I$ is given by the Biot–Savart law:

$$d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{dl \times \hat{r}}{r^2} \quad \text{or} \quad d\mathbf{B} = \frac{\mu_0 I}{4\pi} \frac{dl \times r}{r^3}. \quad (6.92)$$

This law is valid for steady currents.

The field due to an infinitely long solenoid is zero outside and has magnitude $B = \mu_0 nI$ inside, where $n$ is the number of turns per unit length. If a sheet of current has current density $J$, then the change in $B$ across the sheet is $\Delta B = \mu_0 J$.

The Lorentz transformations give the relations between the $\mathbf{E}$ and $\mathbf{B}$ fields in two different frames:

$$ E'_\parallel = E_\parallel, \quad E'_\perp = \gamma (E_\perp + \mathbf{v} \times \mathbf{B}_\perp), $$

$$ B'_\parallel = B_\parallel, \quad B'_\perp = \gamma (B_\perp - (v/c^2) \times E_\perp), \quad (6.93) $$

where $\mathbf{v}$ is the velocity of frame $F'$ with respect to frame $F$. If there exists a frame in which $\mathbf{B} = 0$ (for example, if all the charges are at rest in one frame), then $\mathbf{B}' = -(v/c^2) \times \mathbf{E}'$ in all frames. Similarly, if there exists a frame in which $\mathbf{E} = 0$ (for example, the frame of a neutral current-carrying wire), then $\mathbf{E}' = \mathbf{v} \times \mathbf{B}'$ in all frames.

Henry Rowland demonstrated that a magnetic field is produced not only by a current in a wire, but also by the overall motion of an electrostatically charged object.

In the Hall effect, an external magnetic field causes the charge carriers in a current-carrying wire to pile up on one side of the wire. This causes a transverse electric field inside the wire given by $E_q = -(\mathbf{J} \times \mathbf{B})/nq$. For most purposes, a current of negative charges moving in one direction acts the same as a current of positive charges moving in the other direction. But the Hall effect can be used to determine the sign of the actual charge carriers.

Problems

6.1 Interstellar dust grain  **

This problem concerns the electrically charged interstellar dust grain that was the subject of Exercise 2.38. Its mass, which was not involved in that problem, may be taken as $10^{-16}$ kg. Suppose it is moving quite freely, with speed $v \ll c$, in a plane perpendicular to the interstellar magnetic field, which in that region has a strength of $3 \cdot 10^{-6}$ gauss. How many years will it take to complete a circular orbit?

6.2 Field from power lines  *

A 50 kV direct-current power line consists of two wire conductors 2 m apart. When this line is transmitting 10 MW of power, how strong is the magnetic field midway between the conductors?
6.3 Repelling wires **
Suppose that the current $I_2$ in Fig. 6.4(b) is equal to $I_1$, but reversed, so that $CD$ is repelled by $GH$. Suppose also that $AB$ and $EF$ lie vertically above $GH$, that the lengths $BC$ and $CD$ are 30 and 15 cm, respectively, and that the conductor $BCDE$, which is 1 mm diameter copper wire as in Fig. 6.4(a), has a weight of 0.08 N/m. In equilibrium the deflection of the hanging frame from the vertical is such that $r = 0.5$ cm. How large is the current? Is the equilibrium stable?

6.4 Vector potential for a wire *
Consider the example in Section 6.3, concerning the vector potential for a long straight wire. Rewrite Eqs. (6.33) and (6.34) in terms of Cartesian coordinates, and verify that $\nabla \times A = B$.

6.5 Vector potential for a finite wire **
(a) Recall the example in Section 6.3 dealing with a thin infinite wire carrying current $I$. We showed that the vector potential $A$ given in Eq. (6.34), or equivalently in Eq. (12.272) in the solution to Problem 6.4, correctly produced the desired magnetic field $B$. However, although they successfully produced $B$, there is something fundamentally wrong with those expressions for $A$. What is it? (The infinities at $r = 0$ and $r = \infty$ are technically fine.)

(b) As mentioned at the end of Section 6.3, if you use Eq. (6.44) to calculate $A$ for an infinite wire, you will obtain an infinite result. Your task here is instead to calculate $A$ for a finite wire of length $2L$ (ignore the return path for the current), at a distance $r$ from the center. You can then find an approximate expression for $A$ for large $L$ (is the issue from part (a) fixed?), and then take the curl to obtain $B$, and then take the $L \to \infty$ limit to obtain the $B$ field for a truly infinite wire.

6.6 Zero divergence of $A$ ***
Show that the vector potential given by Eq. (6.44) satisfies $\nabla \cdot A = 0$, provided that the current is steady (that is, $\nabla \cdot J = 0$). Hints: Use the divergence theorem, and be careful about the two types of coordinates in Eq. (6.44) (the 1’s and 2’s). You will need to show that $\nabla_1(1/r_{12}) = -\nabla_2(1/r_{12})$, where the subscript denotes the set of coordinates with respect to which the derivative is taken. The vector identity $\nabla \cdot (fF) = f\nabla \cdot F + F \cdot \nabla f$ will come in handy.

6.7 Vector potential on a spinning sphere ****
A spherical shell with radius $R$ and uniform surface charge density $\sigma$ rotates with angular speed $\omega$ around the $z$ axis. Calculate the vector potential at a point on the surface of the sphere. Do this in three steps as follows.
(a) By direct integration, calculate $A$ at the point $(R, 0, 0)$. You will want to slice the shell into rings whose points are equidistant from $(R, 0, 0)$. The calculation isn’t so bad once you realize that only one component of the velocity survives.

(b) Find $A$ at the point $(x, 0, z)$ in Fig. 6.34 by considering the setup to be the superposition of two shells rotating with the angular velocity vectors $\omega_1$ and $\omega_1$ shown. (This works because angular velocity vectors simply add.)

(c) Finally, determine $A$ at a general point $(x, y, z)$ on the surface of the sphere.

6.8 *The field from a loopy wire*  
A current $I$ runs along an arbitrarily shaped wire that connects two given points, as shown in Fig. 6.35 (it need not lie in a plane). Show that the magnetic field at distant locations is essentially the same as the field due to a straight wire with current $I$ running between the two points.

6.9 *Scaled-up ring*  
Consider two circular rings of copper wire. One ring is a scaled-up version of the other, twice as large in all regards (radius, cross-sectional radius). If currents around the rings are driven by equal voltage sources, how do the magnetic fields at the centers compare?

6.10 **Rings with opposite currents**  
Two parallel rings have the same axis and are separated by a small distance $\epsilon$. They have the same radius $a$, and they carry the same current $I$ but in opposite directions. Consider the magnetic field at points on the axis of the rings. The field is zero midway between the rings, because the contributions from the rings cancel. And the field is zero very far away. So it must reach a maximum value at some point in between. Find this point. Work in the approximation where $\epsilon \ll a$.

6.11 **Field at the center of a sphere**  
A spherical shell with radius $R$ and uniform surface charge density $\sigma$ spins with angular frequency $\omega$ around a diameter. Find the magnetic field at the center.

6.12 **Field in the plane of a ring**  
A ring with radius $R$ carries a current $I$. Show that the magnetic field due to the ring, at a point in the plane of the ring, a distance $a$ from the center (either inside or outside the ring), is given by

$$B = 2 \cdot \frac{\mu_0 I}{4\pi} \int_0^{\frac{\pi}{2}} \frac{(R - a \cos \theta)R d\theta}{(a^2 + R^2 - 2aR \cos \theta)^{3/2}}.$$  (6.94)
Hint: The easiest way to handle the cross product in the Biot–Savart law is to write the Cartesian coordinates of $dl$ and $r$ in terms of an angle $\theta$ in the ring.

This integral can’t be evaluated in closed form (except in terms of elliptic functions), but it can always be evaluated numerically if desired. For the special case of $a = 0$ at the center of the ring, the integral is easy to do; verify that it yields the result given in Eq. (6.54).

6.13 Magnetic dipole **
Consider the result from Problem 6.12. In the $a \gg R$ limit (that is, very far from the ring), make suitable approximations and show that the magnitude of the magnetic field in the plane of the ring is approximately equal to $(\mu_0/4\pi)(m/a^3)$, where $m \equiv \pi R^2 I = (\text{area})I$ is the magnetic dipole moment of the ring. This is a special case of a result we will derive in Chapter 11.

6.14 Far field from a square loop ***
Consider a square loop with current $I$ and side length $a$. The goal of this problem is to determine the magnetic field at a point a large distance $r$ (with $r \gg a$) from the loop.

(a) At the distant point $P$ in Fig. 6.36, the two vertical sides give essentially zero Biot–Savart contributions to the field, because they are essentially parallel to the radius vector to $P$. What are the Biot–Savart contributions from the two horizontal sides? These are easy to calculate because every little interval in these sides is essentially perpendicular to the radius vector to $P$. Show that the sum (or difference) of these contributions equals $\mu_0 I a^2 / 2\pi r^3$, to leading order in $a$.

(b) This result of $\mu_0 I a^2 / 2\pi r^3$ is not the correct field from the loop at point $P$. The correct field is half of this, or $\mu_0 I a^2 / 4\pi r^3$. We will eventually derive this in Chapter 11, where we will show that the general result is $\mu_0 I A / 4\pi r^3$, where $A$ is the area of a loop with arbitrary shape. But we should be able to calculate it via the Biot–Savart law. Where is the error in the reasoning in part (a), and how do you go about fixing it? This is a nice one – don’t peek at the answer too soon!

6.15 Magnetic scalar “potential” **
(a) Consider an infinite straight wire carrying current $I$. We know that the magnetic field outside the wire is $B = (\mu_0 I / 2\pi r) \hat{\theta}$. There are no currents outside the wire, so $\nabla \times B = 0$; verify this by explicitly calculating the curl.

(b) Since $\nabla \times B = 0$, we should be able to write $B$ as the gradient of a function, $B = \nabla \psi$. Find $\psi$, but then explain why the usefulness of $\psi$ as a potential function is limited.
6.16 *Copper solenoid* **
A solenoid is made by winding two layers of No. 14 copper wire on a cylindrical form 8 cm in diameter. There are four turns per centimeter in each layer, and the length of the solenoid is 32 cm. From the wire tables we find that No. 14 copper wire, which has a diameter of 0.163 cm, has a resistance of 0.010 ohm/m at 75°C. (The coil will run hot!) If the solenoid is connected to a 50 V generator, what will be the magnetic field strength at the center of the solenoid in gauss, and what is the power dissipation in watts?

6.17 *A rotating solid cylinder* **
(a) A very long cylinder with radius \( R \) and uniform volume charge density \( \rho \) spins with frequency \( \omega \) around its axis. What is the magnetic field at a point on the axis?
(b) How would your answer change if all the charge were concentrated on the surface?

6.18 *Vector potential for a solenoid* **
A solenoid has radius \( R \), current \( I \), and \( n \) turns per unit length. Given that the magnetic field is \( B = \mu_0 n I \) inside and \( B = 0 \) outside, find the vector potential \( A \) both inside and outside. Do this in two ways as follows.
(a) Use the result from Exercise 6.41.
(b) Use the expression for the curl in cylindrical coordinates given in Appendix F to find the forms of \( A \) that yield the correct values of \( B = \nabla \times A \) in the two regions.

6.19 *Solenoid field, inside and outside* ***
Consider an infinite solenoid with circular cross section. The current is \( I \), and there are \( n \) turns per unit length. Show that the magnetic field is zero outside and \( B = \mu_0 n I \) (in the longitudinal direction) everywhere inside. Do this in three steps as follows.
(a) Show that the field has only a longitudinal component. *Hint:* Consider the contributions to the field from rings that are symmetrically located with respect to a given point.
(b) Use Ampère’s law to show that the field has a uniform value outside and a uniform value inside, and that these two values differ by \( \mu_0 n I \).
(c) Show that \( B \to 0 \) as \( r \to \infty \). There are various ways to do this. One is to obtain an upper bound on the field contribution due to a given ring by unwrapping the ring into a straight wire segment, and then finding the field due to this straight segment.

6.20 *A slab and a sheet* **
A volume current density \( \mathbf{J} = J \hat{z} \) exists in a slab between the infinite planes at \( x = -b \) and \( x = b \). (So the current is coming out
of the page in Fig. 6.37.) Additionally, a surface current density \( J = 2bJ \) points in the \(-\hat{z}\) direction on the plane at \( x = b \).

(a) Find the magnetic field as a function of \( x \), both inside and outside the slab.

(b) Verify that \( \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \) inside the slab. (Don’t worry about the boundaries.)

6.21 Maximum field in a cyclotron **

For some purposes it is useful to accelerate negative hydrogen ions in a cyclotron. A negative hydrogen ion, \( \text{H}^- \), is a hydrogen atom to which an extra electron has become attached. The attachment is fairly weak; an electric field of only \( 4.5 \cdot 10^8 \) \( \text{V/m} \) in the frame of the ion (a rather small field by atomic standards) will pull an electron loose, leaving a hydrogen atom. If we want to accelerate \( \text{H}^- \) ions up to a kinetic energy of 1 GeV \( (10^9 \text{ eV}) \), what is the highest magnetic field we dare use to keep them on a circular orbit up to final energy? (To find \( \gamma \) for this problem you only need the rest energy of the \( \text{H}^- \) ion, which is of course practically the same as that of the proton, approximately 1 GeV.)

6.22 Zero force in any frame **

A neutral wire carries current \( I \). A stationary charge is nearby. There is no electric field from the neutral wire, so the electric force on the charge is zero. And although there is a magnetic field, the charge isn’t moving, so the magnetic force is also zero. The total force on the charge is therefore zero. Hence it must be zero in every other frame. Verify this, in a particular case, by using the Lorentz transformations to find the \( \mathbf{E} \) and \( \mathbf{B} \) fields in a frame moving parallel to the wire with velocity \( \mathbf{v} \).

6.23 No magnetic shield **

A student said, “You almost convinced me that the force between currents, which I thought was magnetism, is explained by electric fields of moving charges. But if so, why doesn’t the metal plate in Fig. 5.1(c) shield one wire from the influence of the other?” Can you explain it?

6.24 \( \mathbf{E} \) and \( \mathbf{B} \) for a point charge **

(a) Use the Lorentz transformations to show that the \( \mathbf{E} \) and \( \mathbf{B} \) fields due to a point charge moving with constant velocity \( \mathbf{v} \) are related by \( \mathbf{B} = (\mathbf{v}/c^2) \times \mathbf{E} \).

(b) If \( v \ll c \), then \( \mathbf{E} \) is essentially obtained from Coulomb’s law, and \( \mathbf{B} \) can be calculated from the Biot–Savart law. Calculate \( \mathbf{B} \) this way, and then verify that it satisfies \( \mathbf{B} = (\mathbf{v}/c^2) \times \mathbf{E} \). (It may be helpful to think of the point charge as a tiny rod of charge, in order to get a handle on the \( dl \) in the Biot–Savart law.)
6.25  **Force in three frames  ***
A charge $q$ moves with speed $v$ parallel to a wire with linear charge density $\lambda$ (as measured in the lab frame). The charges in the wire move with speed $u$ in the opposite direction, as shown in Fig. 6.38. If the charge $q$ is a distance $r$ from the wire, find the force on it in (a) the given lab frame, (b) its own rest frame, (c) the rest frame of the charges in the wire. Do this by calculating the electric and magnetic forces in the various frames. Then check that the force in the charge’s rest frame relates properly to the forces in the other two frames. You can use the fact that the $\gamma$ factor associated with the relativistic addition of $u$ and $v$ is $\gamma_u \gamma_v (1 + \beta_u \beta_v)$.

6.26  **Motion in E and B fields  ***
The task of Exercise 6.29 is to show that if a charged particle moves in the $xy$ plane in the presence of a uniform magnetic field in the $z$ direction, the path will be a circle. What does the path look like if we add on a uniform electric field in the $y$ direction? Let the particle have mass $m$ and charge $q$. And let the magnitudes of the electric and magnetic fields be $E$ and $B$. Assume that the velocity is nonrelativistic, so that $\gamma \approx 1$ (this assumption isn’t necessary in Exercise 6.29, because $v$ is constant there). Be careful, the answer is a bit counterintuitive.

6.27  **Special cases of Lorentz transformations  ***
Figure 6.39 shows four setups involving two infinite charged sheets in a given frame $F$. Another frame $F'$ moves to the right with speed $v$. Explain why these setups demonstrate the six indicated special cases (depending on which field is set equal to zero) of the Lorentz transformations in Eq. (6.76). (Note: one of the sheets has been drawn shorter to indicate length contraction. This is purely symbolic; all of the sheets have infinite length.)

6.28  **The retarded potential  ****
A point charge $q$ moves with speed $v$ along the line $y = r$ in the $xy$ plane. We want to find the magnetic field at the origin at the moment the charge crosses the $y$ axis.

(a) Starting with the electric field in the charge’s frame, use the Lorentz transformation to show that, in the lab frame, the magnitude of the magnetic field at the origin (at the moment the charge crosses the $y$ axis) equals $B = (\mu_0/4\pi)(qv/r^2)$.

(b) Use the Biot–Savart law to calculate the magnetic field at the origin. For the purposes of obtaining the current, you may assume that the “point” charge takes the shape of a very short stick. You should obtain an incorrect answer, lacking the $\gamma$ factor in the above correct answer.
The magnetic field

(c) The Biot–Savart method is invalid because the Biot–Savart law holds for steady currents (or slowly changing ones, but see Footnote 8). But the current due to the point charge is certainly not steady. At a given location along the line of the charge’s motion, the current is zero, then nonzero, then zero again.

For non-steady currents, the validity of the Biot–Savart law can be restored if we use the so-called “retarded time.”\(^{11}\) The basic idea with the retarded time is that, since information can travel no faster than the speed of light, the magnetic field at the origin, at the moment the charge crosses the \(y\) axis, must be related to what the charge was doing at an earlier time. More precisely, this earlier time (the “retarded time”) is the time such that if a light signal were emitted from the charge at this time, then it would reach the origin at the same instant the charge crosses the \(y\) axis. Said in another way, if someone standing at the origin takes a photograph of the surroundings at the moment the charge crosses the \(y\) axis, then the position of the charge in the photograph (which will \textit{not} be on the \(y\) axis) is the charge’s location we are concerned with.\(^{12}\)

\(^{11}\) However, there is an additional term in the modified Biot–Savart law, which makes things more complicated. So we’ll work instead with the vector potential \(\mathbf{A}\), which still has only one term in its modified form. We can then obtain \(\mathbf{B}\) by taking the curl of \(\mathbf{A}\).

\(^{12}\) Due to the finite speed of light, this is quite believable, so we will just accept it as true. But intuitive motivations aside, the modified retarded-time forms of \(\mathbf{B}\) and \(\mathbf{A}\) can be
Your tasks are to: find the location of the charge in the photograph; explain why the length of the little stick representing the charge has a greater length in the photograph than you might naively think; find this length. For the purposes of calculating the vector potential $A$ at the origin, we therefore see that *the current extends over a greater length* than in the incorrect calculation above in part (b). Show that this effect produces the necessary extra $\gamma$ factor in $A$, and hence also in $B$. (Having taken into account the retarded time, the expression for $A$ in Eq. (6.46) remains valid.)

Exercises

6.29 *Motion in a B field*  **
A particle of charge $q$ and rest mass $m$ is moving with velocity $v$ where the magnetic field is $B$. Here $B$ is perpendicular to $v$, and there is no electric field. Show that the path of the particle is a curve with radius of curvature $R$ given by $R = p/qB$, where $p$ is the momentum of the particle, $\gamma mv$. (*Hint:* Note that the force $qv \times B$ can only change the direction of the momentum, not the magnitude. By what angle $\Delta \theta$ is the direction of $p$ changed in a short time $\Delta t$?) If $B$ is the same everywhere, the particle will follow a circular path. Find the time required to complete one revolution.

6.30 *Proton in space*  *
A proton with kinetic energy $10^{16}$ eV ($\gamma = 10^7$) is moving perpendicular to the interstellar magnetic field, which in that region of the galaxy has a strength $3 \cdot 10^{-6}$ gauss. What is the radius of curvature of its path and how long does it take to complete one revolution? (Use the results from Exercise 6.29.)

6.31 *Field from three wires*  *
Three long straight parallel wires are located as shown in Fig. 6.40. One wire carries current $2I$ into the paper; each of the others carries current $I$ in the opposite direction. What is the strength of the magnetic field at the point $P_1$ and at the point $P_2$?

6.32 *Oersted’s experiment*  *
Describing the experiment in which he discovered the influence of an electric current on a nearby compass needle, H. C. Oersted wrote: “If the distance of the connecting wire does not exceed three-quarters of an inch from the needle, the declination of the rigorously derived from Maxwell’s equations, as must be the case for any true statement about electromagnetic fields.

13 If you want to solve this problem by working with the modified Biot–Savart law, it’s a bit trickier. You’ll need to use Eq. (14) in the article mentioned in Footnote 8. And you’ll need to be very careful with all of the lengths involved.
The magnetic field

The magnetic field needle makes an angle of about 45°. If the distance is increased the angle diminishes proportionally. The declination likewise varies with the power of the battery.” About how large a current must have been flowing in Oersted’s “connecting wire”? Assume the horizontal component of the earth’s field in Copenhagen in 1820 was the same as it is today, 0.2 gauss.

6.33 Force between wires *

Suppose the current $I$ that flows in the circuit in Fig. 5.1(b) is 20 amperes. The distance between the wires is 5 cm. How large is the force, per meter of length, that pushes horizontally on one of the wires?

6.34 Torque on a loop ***

The main goal of this problem is to find the torque that acts on a planar current loop in a uniform magnetic field. The uniform field $\mathbf{B}$ points in some direction in space. We shall orient our coordinates so that $\mathbf{B}$ is perpendicular to the $x$ axis, and our current loop lies in the $xy$ plane, as shown in Fig. 6.41. (You should convince yourself that this is always possible.) The shape and size of the (planar) loop are arbitrary; we may think of the current as being supplied by twisted leads on which any net force will be zero. Consider some small element of the loop, and work out its contribution to the torque about the $x$ axis. Only the $z$ component of the force on it will be involved, and hence only the $y$ component of the field $\mathbf{B}$, which we have indicated as $\hat{y}B_y$ in the diagram. Set up the integral that will give the total torque. Show that this integral will give, except for constant factors, the area of the loop.

The magnetic moment of a current loop is defined as a vector $\mathbf{m}$ whose magnitude is $Ia$, where $I$ is the current and $a$ is the area of the loop, and whose direction is normal to the loop with a right-hand-thread relation to the current, as shown in the figure. (We will meet the current loop and its magnetic moment again in Chapter 11.) Show now that your result implies that the torque $\mathbf{N}$ on any current loop is given by the vector equation

$$\mathbf{N} = \mathbf{m} \times \mathbf{B}.$$  \hspace{1cm} (6.95)

What about the net force on the loop?

6.35 Determining $c$ ****

The value of $1/\sqrt{\mu_0\varepsilon_0}$ (or equivalently the value of $c$) can be determined by electrical experiments involving low-frequency fields only. Consider the arrangement shown in Fig. 6.42. The force between capacitor plates is balanced against the force between parallel wires carrying current in the same direction. A voltage alternating sinusoidally at a frequency $f$ (in cycles per second) is applied to the parallel-plate capacitor $C_1$ and also to the capacitor $C_2$. 

Figure 6.41.
The charge flowing into and out of $C_2$ constitutes the current in the rings.

Suppose that $C_2$ and the various distances involved have been adjusted so that the time-average downward force on the upper plate of $C_1$ exactly balances the time-average downward force on the upper ring. (Of course, the weights of the two sides should be adjusted to balance with the voltage turned off.) Show that under these conditions the constant $1/\sqrt{\mu_0\varepsilon_0}$ ($= c$) can be computed from measured quantities as follows:

$$
\frac{1}{\sqrt{\mu_0\varepsilon_0}} = (2\pi)^{3/2}a \left(\frac{b}{h}\right)^{1/2} \left(\frac{C_2}{C_1}\right)f. \tag{6.96}
$$

(If you work with Gaussian instead of SI units, you will end up solving for $c$ instead of $1/\sqrt{\mu_0\varepsilon_0}$.) Assume $s \ll a$ and $h \ll b$.

Note that only measurements of distance and time (or frequency) are required, apart from a measurement of the ratio of the two capacitances $C_1$ and $C_2$. Electrical units, as such, are not involved in the result. (The experiment is actually feasible at a frequency as low as 60 cycles/second if $C_2$ is made, say, $10^6$ times $C_1$ and the current rings are made with several turns to multiply the effect of a small current.)

6.36 *Field at different radii*

A current of 8000 amperes flows through an aluminum rod 4 cm in diameter. Assuming the current density is uniform through the
cross section, find the strength of the magnetic field at 1 cm, at 2 cm, and at 3 cm from the axis of the rod.

6.37 **Off-center hole**

A long copper rod 8 cm in diameter has an off-center cylindrical hole, as shown in Fig. 6.43, down its full length. This conductor carries a current of 900 amp flowing in the direction “into the paper.” What is the direction, and strength in gauss, of the magnetic field at the point $P$ that lies on the axis of the outer cylinder?

6.38 **Uniform field in off-center hole**

A cylindrical rod with radius $R$ carries current $I$ (with uniform current density), with its axis lying along the $z$ axis. A cylindrical cavity with an arbitrary radius is hollowed out from the rod at an arbitrary location; a cross section is shown in Fig. 6.44. Assume that the current density in the remaining part stays the same (which would be the case for a fixed voltage difference between the ends). Let $a$ be the position of the center of the cavity with respect to the center of the rod. Show that the magnetic field inside the cylindrical cavity is uniform (in both magnitude and direction). Hint: Show that the field inside a solid cylinder can be written in the form $\mathbf{B} = \left(\frac{\mu_0 I}{2\pi R^2}\right)\hat{z} \times \mathbf{r}$, and then use superposition with another appropriately chosen cylinder.

6.39 **Constant magnitude of $B$**

How should the current density inside a thick cylindrical wire depend on $r$ so that the magnetic field has constant magnitude inside the wire?

6.40 **The pinch effect**

Since parallel current filaments attract one another, one might think that a current flowing in a solid rod like the conductor in Problem 6.36 would tend to concentrate near the axis of the rod. That is, the conduction electrons, instead of distributing themselves evenly as usual over the interior of the metal, would crowd in toward the axis and most of the current would be there. What do you think prevents this from happening? Ought it happen to any extent at all? Can you suggest an experiment to detect such an effect, if it should exist?

6.41 **Integral of $A$, flux of $B$**

Show that the line integral of the vector potential $\mathbf{A}$ around a closed curve $C$ equals the magnetic flux $\Phi$ through a surface $S$ bounded by the curve. This result is very similar to Ampère’s law, which says that the line integral of the magnetic field $\mathbf{B}$ around a closed curve $C$ equals (up to a factor of $\mu_0$) the current flux $I$ through a surface $S$ bounded by the curve.
6.42 *Finding the vector potential*  
See if you can devise a vector potential that will correspond to a uniform field in the $z$ direction: $B_x = 0, B_y = 0, B_z = B_0$.

6.43 **Vector potential inside a wire**  
A round wire of radius $r_0$ carries a current $I$ distributed uniformly over the cross section of the wire. Let the axis of the wire be the $z$ axis, with $\hat{z}$ the direction of the current. Show that a vector potential of the form $A = A_0\hat{z}(x^2 + y^2)$ will correctly give the magnetic field $B$ of this current at all points inside the wire. What is the value of the constant, $A_0$?

6.44 **Line integral along the axis**  
Consider the magnetic field of a circular current ring, at points on the axis of the ring, given by Eq. (6.53). Calculate explicitly the line integral of the field along the axis from $-\infty$ to $\infty$, to check the general formula

$$\int \mathbf{B} \cdot d\mathbf{s} = \mu_0 I. \quad (6.97)$$

Why may we ignore the “return” part of the path which would be necessary to complete a closed loop?

6.45 **Field from an infinite wire**  
Use the Biot–Savart law to calculate the magnetic field at a distance $b$ from an infinite straight wire carrying current $I$.

6.46 **Field from a wire frame**  
(a) Current $I$ flows around the wire frame in Fig. 6.45(a). What is the direction of the magnetic field at $P$, the center of the cube?  
(b) Show by using superposition that the field at $P$ is the same as if the frame were replaced by the single square loop shown in Fig. 6.45(b).

6.47 **Field at the center of an orbit**  
An electron is moving at a speed $0.01c$ on a circular orbit of radius 10$^{-10}$ m. What is the strength of the resulting magnetic field at the center of the orbit? (The numbers given are typical, in order of magnitude, for an electron in an atom.)

6.48 **Fields from two rings**  
A ring with radius $r$ and linear charge density $\lambda$ spins with frequency $\omega$. A second ring with radius $2r$ has the same density $\lambda$ and frequency $\omega$. Each ring produces a magnetic field at its center. How do the magnitudes of these fields compare?

6.49 **Field at the center of a disk**  
A disk with radius $R$ and surface charge density $\sigma$ spins with angular frequency $\omega$. What is the magnetic field at the center?
6.50 *Hairpin field*
A long wire is bent into the hairpin-like shape shown in Fig. 6.46. Find an exact expression for the magnetic field at the point \( P \) that lies at the center of the half-circle.

6.51 *Current in the earth*
The earth’s metallic core extends out to 3000 km, about half the earth’s radius. Imagine that the field we observe at the earth’s surface, which has a strength of roughly 0.5 gauss at the north magnetic pole, is caused by a current flow in a ring around the “equator” of the core. How big would that current be?

6.52 **Right-angled wire**
A wire carrying current \( I \) runs down the \( y \) axis to the origin, thence out to infinity along the positive \( x \) axis. Show that the magnetic field at any point in the \( xy \) plane (except right on one of the axes) is given by

\[
B_z = \frac{\mu_0 I}{4\pi} \left( \frac{1}{x} + \frac{1}{y} + \frac{x}{y\sqrt{x^2 + y^2}} + \frac{y}{x\sqrt{x^2 + y^2}} \right). \tag{6.98}
\]

6.53 **Superposing right angles**
Use the result from Exercise 6.52, along with superposition, to derive the magnetic field due to an infinite straight wire. (This is certainly an inefficient way of obtaining this field!)

6.54 **Force between a wire and a loop**
Figure 6.47 shows a horizontal infinite straight wire with current \( I_1 \) pointing into the page, passing a height \( z \) above a square horizontal loop with side length \( \ell \) and current \( I_2 \). Two of the sides of the square are parallel to the wire. As with a circular ring, this square produces a magnetic field that points upward on its axis. The field fans out away from the axis. From the right-hand rule, you can show that the magnetic force on the straight wire points to the right. By Newton’s third law, the magnetic force on the square must therefore point to the left.

Your tasks: explain qualitatively, by drawing the fields and forces, why the force on the square does indeed point to the left; then show that the net force equals \( \mu_0 I_1 I_2 \ell^2 / 2\pi R^2 \), where \( R = \sqrt{z^2 + (\ell/2)^2} \) is the distance from the wire to the right and left sides of the square. (The calculation of the force on the wire is a bit more involved. We’ll save that for Exercise 11.20, after we’ve discussed magnetic dipoles.)

6.55 **Helmholtz coils**
One way to produce a very uniform magnetic field is to use a very long solenoid and work only in the middle section of its interior. This is often inconvenient, wasteful of space and power. Can you
suggest ways in which two short coils or current rings might be arranged to achieve good uniformity over a limited region? 

_Hint:_ Consider two coaxial current rings of radius \(a\), separated axially by a distance \(b\). Investigate the uniformity of the field in the vicinity of the point on the axis midway between the two coils. Determine the magnitude of the coil separation \(b\) that for given coil radius \(a\) will make the field in this region as nearly uniform as possible.

6.56 **Field at the tip of a cone**

A hollow cone (like a party hat) has vertex angle \(2\theta\), slant height \(L\), and surface charge density \(\sigma\). It spins around its symmetry axis with angular frequency \(\omega\). What is the magnetic field at the tip?

6.57 *A rotating cylinder*

An infinite cylinder with radius \(R\) and surface charge density \(\sigma\) spins around its symmetry axis with angular frequency \(\omega\). Find the magnetic field inside the cylinder.

6.58 **Rotating cylinders**

Two long coaxial aluminum cylinders are charged to a potential difference of 15 kV. The inner cylinder (assumed to be the positive one) has an outer diameter of 6 cm, the outer cylinder an inner diameter of 8 cm. With the outer cylinder stationary the inner cylinder is rotated around its axis at a constant speed of 30 revolutions per second. Describe the magnetic field this produces and determine its intensity in gauss. What if both cylinders are rotated in the same direction at 30 revolutions per second?

6.59 **Scaled-down solenoid**

Consider two solenoids, one of which is a tenth-scale model of the other. The larger solenoid is 2 meters long, 1 meter in diameter, and is wound with 1 cm diameter copper wire. When the coil is connected to a 120 V direct-current generator, the magnetic field at its center is 1000 gauss. The scaled-down model is exactly one-tenth the size in every linear dimension, including the diameter of the wire. The number of turns is the same, and it is designed to provide the same central field.

(a) Show that the voltage required is the same, namely 120 V.
(b) Compare the coils with respect to the power dissipated and the difficulty of removing this heat by some cooling means.

6.60 **Zero field outside a solenoid**

We showed in the solution to Problem 6.19 that the magnetic field is zero outside an infinite solenoid with arbitrary (uniform) cross-sectional shape. We can demonstrate this fact in another way, similar in spirit to Problem 1.17.
Consider a thin cone emanating from an exterior point $P$, and look at the two patches where it intersects the solenoid. Consider also the thin cone symmetrically located on the other side of $P$ (as shown in Fig. 6.48), along with its two associated patches. Show that the sum of the field contributions due to these four patches is zero at $P$.

6.61 **Rectangular torus**

A coil is wound evenly on a torus of rectangular cross section. There are $N$ turns of wire in all. Only a few are shown in Fig. 6.49. With so many turns, we shall assume that the current on the surface of the torus flows exactly radially on the annular end faces, and exactly longitudinally on the inner and outer cylindrical surfaces. First convince yourself that on this assumption, symmetry requires that the magnetic field everywhere should point in a “circumferential” direction, that is, that all field lines are circles about the axis of the torus. Second, prove that the field is zero at all points outside the torus, including the interior of the central hole. Third, find the magnitude of the field inside the torus, as a function of radius.

6.62 **Creating a uniform field**

For a delicate magnetic experiment, a physicist wants to cancel the earth’s field over a volume roughly $30 \times 30 \times 30$ cm in size, so that the residual field in this region will not be greater than 10 milligauss at any point. The strength of the earth’s field in this location is 0.55 gauss, making an angle of $30^\circ$ with the vertical. It may be assumed constant to a milligauss or so over the volume in question. (The earth’s field itself would hardly vary that much over a foot or so, but in a laboratory there are often local perturbations.) Determine roughly what solenoid dimensions would be suitable for the task, and estimate the number of ampere turns (that is, the current $I$ multiplied by the number of turns $N$) required in your compensating system.

6.63 **Solenoids and superposition**

A number of simple facts about the fields of solenoids can be found by using superposition. The idea is that two solenoids of the same diameter, and length $L$, if joined end to end, make a solenoid of length $2L$. Two semi-infinite solenoids butted together make an infinite solenoid, and so on. (A semi-infinite solenoid is one that has one end here and the other infinitely far away.) Here are some facts you can prove this way.

(a) In the finite-length solenoid shown in Fig. 6.50(a), the magnetic field on the axis at the point $P_2$ at one end is approximately half the field at the point $P_1$ in the center. (Is it slightly more than half, or slightly less than half?)
(b) In the semi-infinite solenoid shown in Fig. 6.50(b), the field line \( FGH \), which passes through the very end of the winding, is a straight line from \( G \) out to infinity.

(c) The flux of \( \mathbf{B} \) through the end face of the semi-infinite solenoid is just half the flux through the coil at a large distance back in the interior.

(d) Any field line that is a distance \( r_0 \) from the axis far back in the interior of the coil exits from the end of the coil at a radius \( r_1 = \sqrt{2} r_0 \), assuming that \( r_0 < \text{(solenoid radius)}/\sqrt{2} \).

Show that these statements are true. What else can you find out?

6.64 *Equal magnitudes* **

Suppose we have a situation in which the component of the magnetic field parallel to the plane of a sheet has the same magnitude on both sides, but changes direction by 90° in going through the sheet. What is going on here? Would there be a force on the sheet? Should our formula for the force on a current sheet apply to cases like this?

6.65 *Proton beam* **

A high-energy accelerator produces a beam of protons with kinetic energy 2 GeV (that is, \( 2 \cdot 10^9 \text{ eV} \) per proton). You may assume that the rest energy of a proton is 1 GeV. The current is 1 milliamp, and the beam diameter is 2 mm. As measured in the laboratory frame:

(a) what is the strength of the electric field caused by the beam 1 cm from the central axis of the beam?

(b) What is the strength of the magnetic field at the same distance?

(c) Now consider a frame \( F' \) that is moving along with the protons. What fields would be measured in \( F' \)?

6.66 *Fields in a new frame* **

In the neighborhood of the origin in the coordinate system \( x, y, z \), there is an electric field \( \mathbf{E} \) of magnitude 100 V/m, pointing in a direction that makes angles of 30° with the \( x \) axis, 60° with the \( y \) axis. The frame \( F' \) has its axes parallel to those just described, but is moving, relative to the first frame, with a speed \( 0.6c \) in the positive \( y \) direction. Find the direction and magnitude of the electric field that will be reported by an observer in the frame \( F' \). What magnetic field does this observer report?

6.67 *Fields from two ions* **

According to observers in the frame \( F \), the following events occurred in the \( xy \) plane. A singly charged positive ion that had been moving with the constant velocity \( v = 0.6c \) in the \( \hat{y} \) direction passed through the origin at \( t = 0 \). At the same instant a similar
ion that had been moving with the same speed, but in the \(-\hat{y}\) direction, passed the point \((2, 0, 0)\) on the \(x\) axis. The distances are in meters.

(a) What are the strength and direction of the electric field, at \(t = 0\), at the point \((3, 0, 0)\)?
(b) What are the strength and direction of the magnetic field at the same place and time?

### 6.68 Force on electrons moving together

Consider two electrons in a cathode ray tube that are moving on parallel paths, side by side, at the same speed \(v\). The distance between them, a distance measured at right angles to their velocity, is \(r\). What is the force that acts on one of them, owing to the presence of the other, as observed in the laboratory frame? If \(v\) were very small compared with \(c\), you could answer \(e^2/4\pi\varepsilon_0 r^2\) and let it go at that. But \(v\) isn’t small, so you have to be careful.

(a) The easiest way to get the answer is as follows. Go to a frame of reference moving with the electrons. In that frame the two electrons are at rest, the distance between them is still \(r\) (why?), and the force is just \(e^2/4\pi\varepsilon_0 r^2\). Now transform the force into the laboratory frame, using the force transformation law, Eq. (5.17). (Be careful about which is the primed system; is the force in the lab frame greater or less than the force in the electron frame?)

(b) It should be possible to get the same answer working entirely in the lab frame. In the lab frame, at the instantaneous position of electron 1, there are both electric and magnetic fields arising from electron 2 (see Fig. 6.30). Calculate the net force on electron 1, which is moving through these fields with speed \(v\), and show that you get the same result as in (a). Make a diagram to show the directions of the fields and forces.

(c) In the light of this, what can you say about the force between two side-by-side moving electrons, in the limit \(v \to c\)?

### 6.69 Relating the forces

Two very long sticks each have uniform linear proper charge density \(\lambda\). (“Proper” means as measured in the rest frame of the given object.) One stick is stationary in the lab frame, while the other stick moves to the left with speed \(v\), as shown in Fig. 6.51. They are \(2r\) apart, and a stationary point charge \(q\) lies midway between them. Find the electric and magnetic forces on the charge \(q\) in the lab frame, and also in the frame of the bottom stick. (Be sure to specify the directions.) Then verify that the total forces in the two frames relate properly.
6.70 **Drifting motion**

Figure 6.52 shows the path of a positive ion moving in the xy plane. There is a uniform magnetic field of 6000 gauss in the z direction. Each period of the ion’s cycloidal motion is completed in 1 microsecond. What are the magnitude and the direction of the electric field that must be present? *Hint:* Think about a frame in which the electric field is zero.

6.71 ***Rowland’s experiment***

Calculate approximately the magnetic field to be expected just above the rotating disk in Rowland’s experiment. Take the relevant data from the description on the page of his paper that is reproduced in Fig. 6.31. You will need to know also that the potential of the rotating disk, with respect to the grounded plates above and below it, was around 10 kilovolts in most of his runs. This information is of course given later in his paper, as is a description of a crucial part of the apparatus, the “astatic” magnetometer shown in the vertical tube on the left. This is an arrangement in which two magnetic needles, oppositely oriented, are rigidly connected together on one suspension so that the torques caused by the earth’s field cancel one another. The field produced by the rotating disk, acting mainly on the nearer needle, can then be detected in the presence of a very much stronger uniform field. That is by no means the only precaution Rowland had to take. In solving this problem, you can make the simplifying assumption that the charges in the disk all travel with their average speed.

6.72 *Transverse Hall field*  
Show that the Gaussian version of Eq. (6.84) must read \( E_t = -J \times B / nqc \), where \( E_t \) is in statvolts/cm, \( B \) is in gauss, \( n \) is in \( \text{cm}^{-3} \), and \( q \) is in esu.

6.73 **Hall voltage**

A Hall probe for measuring magnetic fields is made from arsenic-doped silicon, which has \( 2 \cdot 10^{21} \) conduction electrons per m\(^3\) and a resistivity of 0.016 ohm-m. The Hall voltage is measured across a ribbon of this \( n \)-type silicon that is 0.2 cm wide, 0.005 cm thick, and 0.5 cm long between thicker ends at which it is connected into a 1 V battery circuit. What voltage will be measured across the 0.2 cm dimension of the ribbon when the probe is inserted into a field of 1 kilogauss?
Overview  In this chapter we study the effects of magnetic fields that change with time. Our main result will be that a changing magnetic field causes an electric field. We begin by using the Lorentz force to calculate the emf around a loop moving through a magnetic field. We then make the observation that this emf can be written in terms of the rate of change of the magnetic flux through the loop. The sign of the induced emf is determined by Lenz’s law. If we shift frames so that the loop is now stationary and the source of the magnetic field is moving, we obtain the same result for the emf in terms of the rate of change of flux, as expected. Faraday’s law of induction states that this result holds independent of the cause of the flux change. For example, it applies to the case in which we turn a dial to decrease the magnetic field while keeping all objects stationary. The differential form of Faraday’s law is one of Maxwell’s equations. Mutual inductance is the effect by which a changing current in one loop causes an emf in another loop. This effect is symmetrical between the two loops, as we will prove. Self-inductance is the effect by which a changing current in a loop causes an emf in itself. The most commonly used object with self-inductance is a solenoid, which we call an inductor, symbolized by $L$. The current in an $RL$ circuit changes in a specific way, as we will discover. The energy stored in an inductor equals $LI^2/2$, which parallels the $CV^2/2$ energy stored in a capacitor. Similarly, the energy density in a magnetic field equals $B^2/2\mu_0$, which parallels the $\epsilon_0E^2/2$ energy density in an electric field.
7.1 Faraday’s discovery
Michael Faraday’s account of the discovery of electromagnetic induction begins as follows:

1. The power which electricity of tension possesses of causing an opposite electrical state in its vicinity has been expressed by the general term Induction; which, as it has been received into scientific language, may also, with propriety, be used in the same general sense to express the power which electrical currents may possess of inducing any particular state upon matter in their immediate neighbourhood, otherwise indifferent. It is with this meaning that I purpose using it in the present paper.

2. Certain effects of the induction of electrical currents have already been recognised and described: as those of magnetization; Ampère’s experiments of bringing a copper disc near to a flat spiral; his repetition with electromagnets of Arago’s extraordinary experiments, and perhaps a few others. Still it appeared unlikely that these could be all the effects which induction by currents could produce; especially as, upon dispensing with iron, almost the whole of them disappear, whilst yet an infinity of bodies, exhibiting definite phenomena of induction with electricity of tension, still remain to be acted upon by the induction of electricity in motion.

3. Further: Whether Ampère’s beautiful theory were adopted, or any other, or whatever reservation were mentally made, still it appeared very extraordinary, that as every electric current was accompanied by a corresponding intensity of magnetic action at right angles to the current, good conductors of electricity, when placed within the sphere of this action, should not have any current induced through them, or some sensible effect produced equivalent in force to such a current.

4. These considerations, with their consequence, the hope of obtaining electricity from ordinary magnetism, have stimulated me at various times to investigate experimentally the inductive effect of electric currents. I lately arrived at positive results; and not only had my hopes fulfilled, but obtained a key which appeared to me to open out a full explanation of Arago’s magnetic phenomena, and also to discover a new state, which may probably have great influence in some of the most important effects of electric currents.

5. These results I purpose describing, not as they were obtained, but in such a manner as to give the most concise view of the whole.

This passage was part of a paper Faraday presented in 1831. It is quoted from his “Experimental Researches in Electricity,” published in London in 1839 (Faraday, 1839). There follows in the paper a description of a dozen or more experiments, through which Faraday brought to light every essential feature of the production of electric effects by magnetic action.

By “electricity of tension” Faraday meant electrostatic charges, and the induction he refers to in the first sentence involves nothing more than we have studied in Chapter 3: that the presence of a charge causes a redistribution of charges on conductors nearby. Faraday’s question was, why does an electric current not cause another current in nearby conductors?
Electromagnetic induction

The production of magnetic fields by electric currents had been thoroughly investigated after Oersted’s discovery in 1820. The familiar laboratory source of these “galvanic” currents was the voltaic battery. The most sensitive detector of such currents was a galvanometer. It consisted of a magnetized needle pivoted like a compass needle or suspended by a weak fiber between two coils of wire. Sometimes another needle, outside the coil but connected rigidly to the first needle, was used to compensate for the influence of the earth’s magnetic field (Fig. 7.1(a)). The sketches in Fig. 7.1(b)–(e) represent a few of Faraday’s induction experiments. You must read his own account, one of the classics of experimental science, to appreciate the resourcefulness with which he pressed the search, the alert and open mind with which he viewed the evidence.

In his early experiments Faraday was puzzled to find that a steady current had no detectable effect on a nearby circuit. He constructed various coils of wire, of which Fig. 7.1(a) shows an example, winding two conductors so that they should lie very close together while still separated by cloth or paper insulation. One conductor would form a circuit with the galvanometer. Through the other he would send a strong current from a battery. There was, disappointingly, no deflection of the galvanometer. But in one of these experiments he noticed a very slight disturbance of the galvanometer when the current was switched on and another when it was switched off. Pursuing this lead, he soon established beyond doubt that currents in other conductors are induced, not by a steady current, but by a changing current. One of Faraday’s brilliant experimental tactics at this stage was to replace his galvanometer, which he realized was not a good detector for a brief pulse of current, by a simple small coil in which he put an unmagnetized steel needle; see Fig. 7.1(b). He found that the needle was left magnetized by the pulse of current induced when the primary current was switched on – and it could be magnetized in the opposite sense by the current pulse induced when the primary circuit was broken.

Here is his own description of another experiment:

In the preceding experiments the wires were placed near to each other, and the contact of the inducing one with the battery made when the inductive effect was required; but as the particular action might be supposed to be exerted only at the moments of making and breaking contact, the induction was produced in another way. Several feet of copper wire were stretched in wide zigzag forms, representing the letter W, on one surface of a broad board; a second wire was stretched in precisely similar forms on a second board, so that when brought near the first, the wires should everywhere touch, except that a sheet of thick paper was interposed. One of these wires was connected with the galvanometer, and the other with a voltaic battery. The first wire was then moved towards the second, and as it approached, the needle was deflected. Being then removed, the needle was deflected in the opposite direction. By first making the wires approach and then recede, simultaneously with the vibrations of the needle, the latter soon became very extensive; but when the wires ceased to move from or towards each other, the galvanometer needle soon came to its usual position.
As the wires approximated, the induced current was in the *contrary* direction to the inducing current. As the wires receded, the induced current was in the *same* direction as the inducing current. When the wires remained stationary, there was no induced current.

In this chapter we study the electromagnetic interaction that Faraday explored in those experiments. From our present viewpoint, induction can be seen as a natural consequence of the force on a charge moving in a magnetic field. In a limited sense, we can derive the induction law from what we already know. In following this course we again depart from the historical order of development, but we do so (borrowing Faraday’s own words from the end of the passage first quoted) “to give the most concise view of the whole.”

### 7.2 Conducting rod moving through a uniform magnetic field

Figure 7.2(a) shows a straight piece of wire, or a slender metal rod, supposed to be moving at constant velocity \( v \) in a direction perpendicular to its length. Pervading the space through which the rod moves there is a uniform magnetic field \( B \), constant in time. This could be supplied by a large solenoid enclosing the entire region of the diagram. The reference frame \( F \) with coordinates \( x, y, z \) is the one in which this solenoid is at rest. In the absence of the rod there is no electric field in that frame, only the uniform magnetic field \( B \).

The rod, being a conductor, contains charged particles that will move if a force is applied to them. Any charged particle that is carried along with the rod, such as the particle of charge \( q \) in Fig. 7.2(b), necessarily moves through the magnetic field \( B \) and therefore experiences a force

\[
f = qv \times B.
\]

With \( B \) and \( v \) directed as shown in Fig. 7.2, the force is in the positive \( x \) direction if \( q \) is a positive charge, and in the opposite direction for the negatively charged electrons that are in fact the mobile charge carriers in most conductors. The consequences will be the same, whether negatives or positives, or both, are mobile.

When the rod is moving at constant speed and things have settled down to a steady state, the force \( f \) given by Eq. (7.1) must be balanced, at every point inside the rod, by an equal and opposite force. This can only arise from an electric field in the rod. The electric field develops in this way: the force \( f \) pushes negative charges toward one end of the rod, leaving the other end positively charged. This goes on until these separated charges themselves cause an electric field \( E \) such that, everywhere in the interior of the rod,

\[
qE = -f.
\]
Then the motion of charge relative to the rod ceases. This charge distribution causes an electric field outside the rod, as well as inside. The field outside looks something like that of separated positive and negative charges, with the difference that the charges are not concentrated entirely at the ends of the rod but are distributed along it. The external field is sketched in Fig. 7.3(a). Figure 7.3(b) is an enlarged view of the positively charged end of the rod, showing the charge distribution on the surface and some field lines both outside and inside the conductor. That is the way things look, at any instant of time, in frame $F$.

Let us observe this system from a frame $F'$ that moves with the rod. Ignoring the rod for the moment, we see in this frame $F'$, indicated in Fig. 7.2(c), a magnetic field $B'$ (not much different from $B$ if $v$ is small, by Eq. (6.76)) together with a uniform electric field, as given by Eq. (6.82),

$$E' = v \times B'.$$

(7.3)

This is valid for any value of $v$. When we add the rod to this system, all we are doing is putting a stationary conducting rod into a uniform electric field. There will be a redistribution of charge on the surface of the rod so as to make the electric field zero inside, as in the case of the metal box of Fig. 3.8, or of any other conductor in an electric field. The presence of the magnetic field $B'$ has no influence on this static charge distribution. Figure 7.4(a) shows some electric field lines in the frame $F'$.

This field is the sum of the uniform field in Eq. (7.3) and the field due to the separated positive and negative charges. In the magnified view of the end of the rod in Fig. 7.4(b), we observe that the electric field inside the rod is zero.

Except for the Lorentz contraction, which is second order in $v/c$, the charge distribution seen at one instant in frame $F$, Fig. 7.3(b), is the same as that seen in $F'$. The electric fields differ because the field in Fig. 7.3 is that of the surface charge distribution alone, while the electric field we see in Fig. 7.4 is the field of the surface charge distribution plus the uniform electric field that exists in that frame of reference. An observer in $F$ says, “Inside the rod there has developed an electric field $E = -v \times B$, exerting a force $qE = -qv \times B$ which just balances the force $qv \times B$ that would otherwise cause any charge $q$ to move along the rod.” An observer in $F'$ says, “Inside the rod there is no electric field, because the redistribution of charge on the rod causes there to be zero net internal field, as usual in a conductor. And although there is a uniform magnetic field here, no force arises from it because no charges are moving.” Each account is correct.

### 7.3 Loop moving through a nonuniform magnetic field

What if we made a rectangular loop of wire, as shown in Fig. 7.5, and moved it at constant speed through the uniform field $B'$? To predict what...
will happen, we need only ask ourselves – adopting the frame $F'$ – what would happen if we put such a loop into a uniform electric field. Obviously two opposite sides of the rectangle would acquire some charge, but that would be all. Suppose, however, that the field $B$ in the frame $F$, though constant in time, is *not uniform* in space. To make this vivid, we show in Fig. 7.6 the field $B$ with a short solenoid as its source. This solenoid, together with the battery that supplies its constant current, is fixed near the origin in the frame $F$. (We said earlier there is no electric field in $F$; if we really use a solenoid of finite resistance to provide the field, there will be an electric field associated with the battery and this circuit. It is irrelevant to our problem and can be ignored. Or we can pack the whole solenoid, with its battery, inside a metal box, making sure the total charge is zero.)

Now, with the loop moving with speed $v$ in the $y$ direction, in the frame $F$, let its position at some instant $t$ be such that the magnetic field strength is $B_1$ at the left side of the loop and $B_2$ along the right side (Fig. 7.6). Let $f$ denote the force that acts on a charge $q$ that rides along with the loop. This force is a function of position on the loop, at this instant of time. Let's evaluate the line integral of $f$, taken around the whole loop (counterclockwise as viewed from above). On the two sides of the loop that lie parallel to the direction of motion, $f$ is perpendicular to the path element $ds$, so these give nothing. Taking account of the contributions from the other two sides, each of length $w$, we have

$$\int f \cdot ds = qv(B_1 - B_2)w. \quad (7.4)$$

If we imagine a charge $q$ to move all around the loop, in a time short enough so that the position of the loop has not changed appreciably, then Eq. (7.4) gives the work done by the force $f$. The work done *per unit charge* is $(1/q) \int f \cdot ds$. We call this quantity *electromotive force*. We use the symbol $\mathcal{E}$ for it, and often shorten the name to emf. So we have

$$\mathcal{E} = \frac{1}{q} \int f \cdot ds \quad (7.5)$$

$\mathcal{E}$ has the same dimensions as electric potential, so the SI unit is the volt, or joule per coulomb. In the Gaussian system, $\mathcal{E}$ is measured in statvolts, or ergs per esu.

We have noted that the force $f$ does work. However, $f$ is a magnetic force, and we know that magnetic forces do no work, because the force is always perpendicular to the velocity. So we seem to have an issue here. Is the magnetic force somehow doing work? If not, then what is? This is the subject of Problem 7.2.

The term *electromotive force* was introduced earlier, in Section 4.9. It was defined as the work per unit charge involved in moving a charge
Electromagnetic induction

around a circuit containing a voltaic cell. We now broaden the definition of emf to include any influence that causes charge to circulate around a closed path. If the path happens to be a physical circuit with resistance \( R \), then the emf \( \mathcal{E} \) will cause a current to flow according to Ohm’s law: \( I = \mathcal{E}/R \). Note that since \( \text{curl } \mathbf{E} = 0 \) for an electrostatic field, such a field cannot cause a charge to circulate around a closed path. By our above definition of electromotive force, an emf must therefore be nonelectrostatic in origin. See Varney and Fisher (1980) for a discussion of electromotive force.

In the particular case we are considering, \( \mathbf{f} \) is the force that acts on a charge moving in a magnetic field, and \( \mathcal{E} \) has the magnitude

\[
\mathcal{E} = \nu w (B_1 - B_2). \tag{7.6}
\]

The electromotive force given by Eq. (7.6) is related in a very simple way to the rate of change of magnetic flux through the loop. (We will be quantitative about this in Theorem 7.1.) By the magnetic flux through a loop we mean the surface integral of \( \mathbf{B} \) over a surface that has the loop for its boundary. The flux \( \Phi \) through the closed curve or loop \( C \) in Fig. 7.7(a) is given by the surface integral of \( \mathbf{B} \) over \( S_1 \):

\[
\Phi_{S_1} = \int_{S_1} \mathbf{B} \cdot d\mathbf{a}_1. \tag{7.7}
\]

We could draw infinitely many surfaces bounded by \( C \). Figure 7.7(b) shows another one, \( S_2 \). Why don’t we have to specify which surface to use in computing the flux? It doesn’t make any difference because \( \int \mathbf{B} \cdot d\mathbf{a} \) will have the same value for all surfaces. Let’s take a minute to settle this point once and for all.

The flux through \( S_2 \) will be \( \int_{S_2} \mathbf{B} \cdot d\mathbf{a}_2 \). Note that we let the vector \( d\mathbf{a}_2 \) stick out from the upper side of \( S_2 \), to be consistent with our choice of side of \( S_1 \). This will give a positive number if the net flux through \( C \) is upward:

\[
\Phi_{S_2} = \int_{S_2} \mathbf{B} \cdot d\mathbf{a}_2. \tag{7.8}
\]

We learned in Section 6.2 that the magnetic field has zero divergence: \( \text{div } \mathbf{B} = 0 \). It follows then from Gauss’s theorem that, if \( S \) is any closed surface (“balloon”) and \( V \) is the volume inside it, we have

\[
\int_S \mathbf{B} \cdot d\mathbf{a} = \int_V \text{div } \mathbf{B} \, dv = 0. \tag{7.9}
\]

Apply this to the closed surface, rather like a kettledrum, formed by joining our \( S_1 \) to \( S_2 \), as in Fig. 7.7(c). On \( S_2 \) the outward normal is opposite the vector \( d\mathbf{a}_2 \) we used in calculating the flux through \( C \). Thus

\[
0 = \int_S \mathbf{B} \cdot d\mathbf{a} = \int_{S_1} \mathbf{B} \cdot d\mathbf{a}_1 + \int_{S_2} \mathbf{B} \cdot (-d\mathbf{a}_2). \tag{7.10}
\]
or
\[ \int_{S_1} \mathbf{B} \cdot d\mathbf{a}_1 = \int_{S_2} \mathbf{B} \cdot d\mathbf{a}_2. \] (7.11)

This shows that it doesn’t matter which surface we use to compute the flux through \( C \).

This is all pretty obvious if you realize that \( \text{div} \mathbf{B} = 0 \) implies a kind of spatial conservation of flux. As much flux enters any volume as leaves it. (We are considering the situation in the whole space at one instant of time.) It is often helpful to visualize “tubes” of flux. A flux tube (Fig. 7.8) is a surface at every point on which the magnetic field line lies in the plane of the surface. It is a surface through which no flux passes, and we can think of it as containing a certain amount of flux, as a fiber optic cable contains fibers. Through any closed curve drawn tightly around a flux tube, the same flux passes. This could be said about the electric field \( \mathbf{E} \) only for regions where there is no electric charge, since \( \text{div} \mathbf{E} = \rho/\varepsilon_0 \).

The magnetic field always has zero divergence everywhere.

Returning now to the moving rectangular loop, let us find the rate of change of flux through the loop. In time \( dt \) the loop moves a distance \( v dt \). This changes in two ways the total flux through the loop, which is \( \int \mathbf{B} \cdot d\mathbf{a} \) over a surface spanning the loop. As you can see in Fig. 7.9, flux is gained at the right, in amount \( B_2 wv dt \), while an amount of flux \( B_1 wv dt \) is lost at the left. Hence \( d\Phi \), the change in flux through the loop in time \( dt \), is

\[ d\Phi = -(B_1 - B_2) w v dt. \] (7.12)
Comparing Eq. (7.12) with Eq. (7.6), we see that, in this case at least, the electromotive force can be expressed as $E = -\frac{d\Phi}{dt}$. It turns out that this is a general result, as the following theorem states.

**Theorem 7.1** If the magnetic field in a given frame is constant in time, then for a loop of any shape moving in any manner, the emf $E$ around the loop is related to the magnetic flux $\Phi$ through the loop by

$$E = -\frac{d\Phi}{dt}. \quad (7.13)$$

**Proof** The loop $C$ in Fig. 7.10 occupies the position $C_1$ at time $t$, and it is moving so that it occupies the position $C_2$ at time $t + dt$. A particular element of the loop $ds$ has been transported with velocity $v$ to its new position. $S$ indicates a surface that spans the loop at time $t$. The flux through the loop at this instant of time is

$$\Phi(t) = \int_S \mathbf{B} \cdot d\mathbf{a}. \quad (7.14)$$

The magnetic field $\mathbf{B}$ comes from sources that are stationary in our frame of reference and remains constant in time, at any point fixed in this frame. At time $t + dt$ a surface that spans the loop is the original surface $S$, left fixed in space, augmented by the “rim” $dS$. (Remember, we are allowed to use any surface spanning the loop to compute the flux through it.) Thus

$$\Phi(t + dt) = \int_{S+dS} \mathbf{B} \cdot d\mathbf{a} = \Phi(t) + \int_{dS} \mathbf{B} \cdot d\mathbf{a}. \quad (7.15)$$

Hence the change in flux, in time $dt$, is just the flux $\int_{dS} \mathbf{B} \cdot d\mathbf{a}$ through the rim $dS$. On the rim, an element of surface area $d\mathbf{a}$ can be expressed as $(v dt) \times ds$, because this cross product has magnitude $|v dt||ds| \sin \theta$ and points in the direction perpendicular to both $v dt$ and $ds$; the $\sin \theta$ in the magnitude gives the correct area of the little parallelogram in Fig. 7.10. So the integral over the surface $dS$ can be written as an integral around the path $C$, in this way:

$$d\Phi = \int_{dS} \mathbf{B} \cdot d\mathbf{a} = \int_C \mathbf{B} \cdot [(v dt) \times ds]. \quad (7.16)$$

Since $dt$ is a constant for the integration, we can factor it out to obtain

$$\frac{d\Phi}{dt} = \int_C \mathbf{B} \cdot (v \times ds). \quad (7.17)$$

The product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ of any three vectors satisfies the relation $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{(b \times a)} \cdot \mathbf{c}$, which you can verify by explicitly writing out
each side in Cartesian components. Using this identity to rearrange the integrand in Eq. (7.17), we have

\[
\frac{d\Phi}{dt} = -\int_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{s}. \tag{7.18}
\]

Now, the force on a charge \( q \) that is carried along by the loop is just \( q\mathbf{v} \times \mathbf{B} \), so the electromotive force, which is the line integral around the loop of the force per unit charge, is just

\[
\mathcal{E} = \int_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{s}. \tag{7.19}
\]

Comparing Eq. (7.18) with Eq. (7.19), we get the simple relation given in Eq. (7.13), valid for arbitrary shape and motion of the loop. (We did not even have to assume that \( \mathbf{v} \) is the same for all parts of the loop!) In summary, the line integral around a moving loop of \( \mathbf{f}/q \), the force per unit charge, is just the negative of the rate of change of flux through the loop.

The sense of the line integral and the direction in which flux is called positive are to be related by a right-hand-thread rule. For instance, in Fig. 7.6, the flux is upward through the loop and is decreasing. Taking the minus sign in Eq. (7.13) into account, our rule would predict an electromotive force that would tend to drive a positive charge around the loop in a counterclockwise direction, as seen looking down on the loop (Fig. 7.11).

There is a better way to look at this question of sign and direction. Note that if a current should flow in the direction of the induced electromotive force, in the situation shown in Fig. 7.11, this current itself would create some flux through the loop in a direction to counteract the assumed flux change (because the Biot–Savart law, Eq. (6.49), tells us that the contributions from this current to the \( \mathbf{B} \) field inside the loop all point upward in Fig. 7.11). That is an essential physical fact, and not the consequence of an arbitrary convention about signs and directions. It is a manifestation of the tendency of systems to resist change. In this context it is traditionally called Lenz’s law.

**Lenz’s law** *The direction of the induced electromotive force is such that the induced current creates a magnetic field that opposes the change in flux.*

Another example of Lenz’s law is illustrated in Fig. 7.12. The conducting ring is falling in the magnetic field of the coil. The flux through the ring is downward and is increasing in magnitude. To counteract this change, some new flux upward is needed. It would take a current flowing around the ring in the direction of the arrows to produce such flux. Lenz’s law assures us that the induced emf will be in the correct direction to cause such a current.
Electromagnetic induction

If the electromotive force causes current to flow in the loop that is shown in Figs. 7.6 and 7.11, as it will if the loop has a finite resistance, some energy will be dissipated in the wire. What supplies this energy? To answer that, consider the force that acts on the current in the loop if it flows in the sense indicated by the arrow in Fig 7.11. The side on the right, in the field $B_2$, will experience a force toward the right, while the opposite side of the loop, in the field $B_1$, will be pushed toward the left. But $B_1$ is greater than $B_2$, so the net force on the loop is toward the left, opposing the motion. To keep the loop moving at constant speed, some external agency has to do work, and the energy thus invested eventually shows up as heat in the wire (see Exercise 7.30). Imagine what would happen if Lenz’s law were violated, or if the force on the loop were to act in a direction to assist the motion of the loop!

Example (Sinusoidal $\mathcal{E}$) A very common element in electrical machinery and electrical instruments is a loop or coil that rotates in a magnetic field. Let’s apply what we have just learned to the system shown in Fig. 7.13, a single loop rotating at constant speed in a magnetic field that is approximately uniform. The mechanical essentials, shaft, bearings, drive, etc., are not drawn. The field $B$ is provided by the two fixed coils. Suppose the loop rotates with angular velocity $\omega$, in radians/second. If its position at any instant is specified by the angle $\theta$, then $\theta = \omega t + \alpha$, where the constant $\alpha$ is simply the position of the loop at $t = 0$. The component of $B$ perpendicular to the plane of the loop is $B \sin \theta$. Therefore the flux through the loop at time $t$ is

$$\Phi(t) = SB \sin(\omega t + \alpha),$$  \hspace{1cm} (7.20)

where $S$ is the area of the loop. For the induced electromotive force we then have

$$\mathcal{E} = - \frac{d\Phi}{dt} = -SB\omega \cos(\omega t + \alpha).$$  \hspace{1cm} (7.21)

If the loop instead of being closed is connected through slip rings to external wires, as shown in Fig. 7.13, we can detect at these terminals a sinusoidally alternating potential difference.

A numerical example will show how the units work out. Suppose the area of the loop in Fig. 7.13 is 80 cm$^2$, the field strength $B$ is 50 gauss, and the loop is rotating at 30 revolutions per second. Then $\omega = 2\pi \cdot 30$, or 188 radians/second. The amplitude, that is, the maximum magnitude of the oscillating electromotive force induced in the loop, is

$$\mathcal{E}_0 = SB\omega = (0.008 \text{ m}^2)(0.005 \text{ tesla})(188 \text{ s}^{-1}) = 7.52 \cdot 10^{-3} \text{ V}. \hspace{1cm} (7.22)$$

You should verify that 1 m$^2 \cdot$tesla/s is indeed equivalent to 1 volt.

7.4 Stationary loop with the field source moving

We can, if we like, look at the events depicted in Fig. 7.6 from a frame of reference that is moving with the loop. That can’t change the physics, only the words we use to describe it. Let $F'$, with coordinates $x', y', z'$,
be the frame attached to the loop, which we now regard as stationary (Fig. 7.14). The coil and battery, stationary in frame $F$, are moving in the $-y'$ direction with velocity $v' = -v$. Let $B'_1$ and $B'_2$ be the magnetic field measured at the two ends of the loop by observers in $F'$ at some instant $t'$. At these positions there will be an electric field in $F'$. Equation (6.82) tells us that

$$E'_1 = v \times B'_1 \quad \text{and} \quad E'_2 = v \times B'_2. \quad (7.23)$$

For observers in $F'$ this is a genuine electric field. It is not an electrostatic field; the line integral of $E'$ around any closed path in $F'$ is not generally zero. In fact, from Eq. (7.23) the line integral of $E'$ around the rectangular loop is

$$\int E' \cdot ds' = wv(B'_1 - B'_2). \quad (7.24)$$

We can call the line integral in Eq. (7.24) the electromotive force $\mathcal{E}'$ on this path. If a charged particle moves once around the path, $\mathcal{E}'$ is the work done on it, per unit charge. $\mathcal{E}'$ is related to the rate of change of flux through the loop. To see this, note that, while the loop itself is stationary, the magnetic field pattern is now moving with the velocity $-v$ of the source. Hence for the flux lost or gained at either end of the loop, in a time interval $dt'$, we get a result similar to Eq. (7.12), and we conclude that

$$\mathcal{E}' = -\frac{d\Phi'}{dt'}. \quad (7.25)$$

Figure 7.12.
As the ring falls, the downward flux through the ring is increasing. Lenz’s law tells us that the induced emf will be in the direction indicated by the arrows, for that is the direction in which current must flow to produce upward flux through the ring. The system reacts so as to oppose the change that is occurring.

Figure 7.13.
The two coils produce a magnetic field $B$ that is approximately uniform in the vicinity of the loop. In the loop, rotating with angular velocity $\omega$, a sinusoidally varying electromotive force is induced.
As observed in the frame $F'$, the loop is at rest and the field source is moving. The fields $\mathbf{B}'$ and $\mathbf{E}'$ are both present and are functions of both position and time.

We can summarize as follows the descriptions in the two frames of reference, $F$, in which the source of $\mathbf{B}$ is at rest, and $F'$, in which the loop is at rest.

- An observer in $F$ says, “We have here a magnetic field that, though it is not uniform spatially, is constant in time. There is no electric field. That wire loop over there is moving with velocity $\mathbf{v}$ through the magnetic field, so the charges in it are acted on by a force $\mathbf{v} \times \mathbf{B}$ per unit charge. The line integral of this force per unit charge, taken around the whole loop, is the electromotive force $\mathbf{E}$, and it is equal to $-d\Phi/dt$. The flux $\Phi$ is $\int \mathbf{B} \cdot d\mathbf{a}$ over a surface $S$ that, at some instant of time $t$ by my clock, spans the loop.”

- An observer in $F'$ says, “This loop is stationary, and only an electric field could cause the charges in it to move. But there is in fact an electric field $\mathbf{E}'$. It seems to be caused by that magnetlike object which happens at this moment to be whizzing by with a velocity $-\mathbf{v}$, producing at the same time a magnetic field $\mathbf{B}'$. The electric field is such that $\int \mathbf{E}' \cdot d\mathbf{s}'$ around this stationary loop is not zero but instead is equal to the negative of the rate of change of flux through the loop, $-d\Phi'/dt'$. The flux $\Phi'$ is $\int \mathbf{B}' \cdot d\mathbf{a}'$ over a surface spanning the loop, the values of $B'$ to be measured all over this surface at some one instant $t'$, by my clock.”

Our conclusions so far are relativistically exact. They hold for any speed $v < c$ provided we observe scrupulously the distinctions between $\mathbf{B}$ and $\mathbf{B}'$, $t$ and $t'$, etc. If $v \ll c$, so that $v^2/c^2$ can be neglected, $\mathbf{B}'$ will be practically equal to $\mathbf{B}$, and we can safely ignore also the distinction between $t$ and $t'$.
### 7.5 Universal law of induction

Let’s carry out three experiments with the apparatus shown in Fig. 7.15. The tables are on wheels so that they can be easily moved. A sensitive galvanometer has been connected to our old rectangular loop, and to increase any induced electromotive force we put several turns of wire in the loop rather than one. Frankly though, our sensitivity might still be marginal, with the feeble source of magnetic field pictured. Perhaps you can devise a more practical version of the experiment.

**Experiment I.** With constant current in the coil and table 1 stationary, table 2 moves toward the right (away from table 1) with speed \( v \). The galvanometer deflects. We are not surprised; we have already analyzed this situation in Section 7.3.

**Experiment II.** With constant current in the coil and table 2 stationary, table 1 moves to the left (away from table 2) with speed \( v \). The galvanometer deflects. This doesn’t surprise us either. We have just discussed in Section 7.4 the equivalence of Experiments I and II, an equivalence that is an example of Lorentz invariance or, for the low speeds of our tables, Galilean invariance. We know that in both experiments the deflection of the galvanometer can be related to the rate of change of flux of \( B \) through the loop.

**Experiment III.** Both tables remain at rest, but we vary the current \( I \) in the coil by sliding the contact \( K \) along the resistance strip. We do this in such a way that the rate of decrease of the field \( B \) at the loop is the same as it was in Experiments I and II. *Does the galvanometer deflect?*

---

**Figure 7.15.**
We imagine that either table can move or, with both tables fixed, the current \( I \) in the coil can be gradually changed.
For an observer stationed at the loop on table 2 and measuring the magnetic field in that neighborhood as a function of time and position, there is no way to distinguish among Experiments I, II, and III. Imagine a black cloth curtain between the two tables. Although there might be minor differences between the field configurations for II and III, an observer who did not know what was behind the curtain could not decide, on the basis of local $B$ measurements alone, which case it was. Therefore if the galvanometer did not respond with the same deflection in Experiment III, it would mean that the relation between the magnetic and electric fields in a region depends on the nature of a remote source. Two magnetic fields essentially similar in their local properties would have associated electric fields with different values of $\int E \cdot ds$.

We find by experiment that III is equivalent to I and II. The galvanometer deflects, by the same amount as before. Faraday’s experiments were the first to demonstrate this fundamental fact. The electromotive force we observe depends only on the rate of change of the flux of $B$, and not on anything else. We can state as a universal relation Faraday’s law of induction:

$$E = \int_C E \cdot ds = -\frac{d}{dt} \int_S B \cdot da = -\frac{d\Phi}{dt} \quad \text{(Faraday’s law)}$$

Using the vector derivative curl, we can express this law in differential form. If the relation

$$\int_C E \cdot ds = -\frac{d}{dt} \int_S B \cdot da$$

is true for any curve $C$ and spanning surface $S$, as our law asserts, it follows that, at any point,

$$\text{curl } E = -\frac{dB}{dt}. \quad (7.28)$$

To show that Eq. (7.28) follows from Eq. (7.27), we proceed as usual to let $C$ shrink down around a point, which we take to be a nonsingular point for the function $B$. Then in the limit the variation of $B$ over the small patch of surface $a$ that spans $C$ will be negligible and the surface integral will approach simply $B \cdot a$. By definition (see Eq. (2.80)), the limit approached by $\int_C E \cdot ds$ as the patch shrinks is $a \cdot \text{curl } E$. Thus Eq. (7.27) becomes, in the limit,

$$a \cdot \text{curl } E = -\frac{d}{dt} (B \cdot a) = a \cdot \left(-\frac{dB}{dt}\right). \quad (7.29)$$
Since this holds for any infinitesimal \( \mathbf{a} \), it must be that\(^1\)

\[
\text{curl} \mathbf{E} = -\frac{dB}{dt}, \tag{7.30}
\]

Recognizing that \( \mathbf{B} \) may depend on position as well as time, we write \( \partial \mathbf{B}/\partial t \) in place of \( dB/dt \). We have then these two entirely equivalent statements of the law of induction:

\[
\begin{align*}
\int_C \mathbf{E} \cdot d\mathbf{s} &= -\frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{a} \\
\text{curl} \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \tag{7.31}
\end{align*}
\]

With Faraday’s law of induction, we are one step closer to the complete set of Maxwell’s equations. We will obtain the last piece to the puzzle in Chapter 9.

In Eq. (7.31) the electric field \( \mathbf{E} \) is to be expressed in our SI units of volts/meter, with \( \mathbf{B} \) in teslas, \( ds \) in meters, and \( da \) in m\(^2\). The electromotive force \( \mathcal{E} = \int_C \mathbf{E} \cdot ds \) will then be given in volts. In Gaussian units the relation expressed by Eq. (7.31) looks like this:

\[
\begin{align*}
\int_C \mathbf{E} \cdot d\mathbf{s} &= -\frac{1}{c} \frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{a} \\
\text{curl} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} \tag{7.32}
\end{align*}
\]

Here \( \mathbf{E} \) is in statvolts/cm, \( \mathbf{B} \) is in gauss, \( ds \) and \( da \) are in cm and cm\(^2\), respectively, and \( c \) is in cm/s. The electromotive force \( \mathcal{E} = \int_C \mathbf{E} \cdot ds \) will be given in statvolts.

The magnetic flux \( \Phi \), which is \( \int_C \mathbf{B} \cdot d\mathbf{a} \), is expressed in tesla-m\(^2\) in our SI units, and in gauss-cm\(^2\), a unit exactly 10\(^8\) times smaller, in Gaussian units (because 1 m\(^2\) = 10\(^4\) cm\(^2\) and 1 tesla = 10\(^4\) gauss, exactly). The SI flux unit is assigned a name of its own, the weber.

When in doubt about the units, you may find one of the following equivalent statements helpful:

- Electromotive force in statvolts equals:
  1/\(c\) times rate of change of flux in gauss-cm\(^2\)/s.
- Electromotive force in volts equals:
  rate of change of flux in tesla-m\(^2\)/s.
- Electromotive force in volts equals:
  10\(^{-8}\) times rate of change of flux in gauss-cm\(^2\)/s.

---

\(^1\) If that isn’t obvious, note that choosing \( \mathbf{a} \) in the \( x \) direction will establish that (curl \( \mathbf{E} \))\(x\) = \( −dB_x/dt \), and so on.
Alternating current in the coils produces a magnetic field which, at the center, oscillates between 50 gauss upward and 50 gauss downward. At any instant the field is approximately uniform within the circle $C$.

If these seem confusing, don’t try to remember them. Just remember that you can look them up on this page.

The differential expression, $\text{curl} E = -\partial B/\partial t$, brings out rather plainly the point we tried to make earlier about the local nature of the field relations. The variation in time of $B$ in a neighborhood completely determines $\text{curl} E$ there – nothing else matters. That does not completely determine $E$ itself, of course. Without affecting this relation, any electrostatic field with $\text{curl} E = 0$ could be superposed.

**Example (Sinusoidal B field)** As a concrete example of Faraday’s law, suppose coils like those in Fig. 7.13 are supplied with 60 cycles per second alternating current, instead of direct current. The current and the magnetic field vary as $\sin(2\pi \cdot 60 \text{ s}^{-1} \cdot t)$, or $\sin(377 \text{ s}^{-1} \cdot t)$. Suppose the amplitude of the current is such that the magnetic field $B$ in the central region reaches a maximum value of 50 gauss, or 0.005 tesla. We want to investigate the induced electric field, and the electromotive force, on the circular path 10 cm in radius shown in Fig. 7.16. We may assume that the field $B$ is practically uniform in the interior of this circle, at any instant of time. So we have

$$B = (0.005 \text{ T}) \sin(377 \text{ s}^{-1} \cdot t). \quad (7.33)$$

The flux through the loop $C$ is

$$\Phi = \pi r^2 B = \pi \cdot (0.1 \text{ m})^2 \cdot (0.005 \text{ T}) \sin(377 \text{ s}^{-1} \cdot t)$$

$$= 1.57 \cdot 10^{-4} \sin(377 \text{ s}^{-1} \cdot t) \text{ T m}^2. \quad (7.34)$$
Using Eq. (7.26) to calculate the electromotive force, we obtain

\[ E = -\frac{d\Phi}{dt} = -(377 \text{ s}^{-1}) \cdot 1.57 \cdot 10^{-4} \cos(377 \text{ s}^{-1} \cdot t) \text{ Tm}^2 \]

\[ = -0.059 \cos(377 \text{ s}^{-1} \cdot t) \text{ V}. \]

The maximum attained by \( E \) is 59 millivolts. The minus sign will ensure that Lenz’s law is respected, if we have defined our directions consistently. The variation of both \( \Phi \) and \( E \) with time is shown in Fig. 7.17.

What about the electric field itself? Usually we cannot deduce \( E \) from a knowledge of \( \text{curl} \, E \) alone. However, our path \( C \) is here a circle around the center of a symmetrical system. If there are no other electric fields around, we may assume that, on the circle \( C \), \( E \) lies in that plane and has a constant magnitude. Then it is a trivial matter to predict its magnitude, since \( \oint_C E \cdot ds = 2\pi rE = E \), which we have already calculated. In this case, the electric field on the circle might look like Fig. 7.18(a) at a particular instant. But if there are other field sources, it could look quite different. If there happened to be a positive and a negative charge located on the axis as shown in Fig. 7.18(b), the electric field in the vicinity of the circle would be the superposition of the electrostatic field of the two charges and the induced electric field.

A consequence of Faraday’s law of induction is that Kirchhoff’s loop rule (which states that \( \oint E \cdot ds = 0 \) around a closed path) is no longer valid in situations where there is a changing magnetic field. Faraday has taken us beyond the comfortable realm of conservative electric fields. The voltage difference between two points now depends on the path between them. Problem 7.4 provides an instructive example of this fact.

A note on the terminology: the term “potential difference” is generally reserved for electrostatic fields, because it is only for such fields that we can uniquely define a potential function \( \phi \) at all points in space, with the property that \( E = -\nabla \phi \). For these fields, the potential difference between points \( a \) and \( b \) is given by \( \phi_b - \phi_a = -\int_a^b E \cdot ds \). The term “voltage difference” applies to any electric field, not necessarily electrostatic, and it is defined similarly as \( V_b - V_a = -\int_a^b E \cdot ds \). If there are changing magnetic fields involved, this line integral will depend on the path between \( a \) and \( b \). The voltage difference is what a voltmeter measures, and we can hook up a voltmeter to any type of circuit, of course, no matter what kinds of electric fields it involves. But if there are changing magnetic fields, Problem 7.4 shows that it matters how we hook it up.

See Romer (1982) for more discussion of this issue.

7.6 Mutual inductance

Two circuits, or loops, \( C_1 \) and \( C_2 \) are fixed in position relative to one another (Fig. 7.19). By some means, such as a battery and a variable resistance, a controllable current \( I_1 \) is caused to flow in circuit \( C_1 \). Let
Let \( \mathbf{B}_1(x, y, z) \) be the magnetic field that would exist if the current in \( C_1 \) remained constant at the value \( I_1 \), and let \( \Phi_{21} \) denote the flux of \( \mathbf{B}_1 \) through the circuit \( C_2 \). Thus

\[
\Phi_{21} = \int_{S_2} \mathbf{B}_1 \cdot d\mathbf{a}_2,
\]

where \( S_2 \) is a surface spanning the loop \( C_2 \). With the shape and relative position of the two circuits fixed, \( \Phi_{21} \) will be proportional to \( I_1 \):

\[
\frac{\Phi_{21}}{I_1} = \text{constant} \equiv M_{21}.
\]

Suppose now that \( I_1 \) changes with time, but *slowly enough* so that the field \( \mathbf{B}_1 \) at any point in the vicinity of \( C_2 \) is related to the current \( I_1 \) in \( C_1 \) (at the same instant of time) in the same way as it would be related for a steady current. (To see why such a restriction is necessary, imagine that \( C_1 \) and \( C_2 \) are 10 meters apart and we cause the current in \( C_1 \) to double in value in 10 nanoseconds!) The flux \( \Phi_{21} \) will change in proportion as \( I_1 \) changes. There will be an electromotive force induced in circuit \( C_2 \), of magnitude

\[
\mathcal{E}_{21} = -\frac{d\Phi_{21}}{dt} \quad \Rightarrow \quad \mathcal{E}_{21} = -M_{21} \frac{dI_1}{dt}.
\]

In Gaussian units there is a factor of \( c \) in the denominator here. But we can define a new constant \( M'_{21} \equiv M_{21}/c \) so that the relation between \( \mathcal{E}_{21} \) and \( dI_1/dt \) remains of the same form.

We call the constant \( M_{21} \) the coefficient of *mutual inductance*. Its value is determined by the geometry of our arrangement of loops. The units will of course depend on our choice of units for \( \mathcal{E}, I, \) and \( t \). In SI

---

**Figure 7.18.**
The electric field on the circular path \( C \). (a) In the absence of sources other than the symmetrical, oscillating current. (b) Including the electrostatic field of two charges on the axis.

**Figure 7.19.**
Current \( I_1 \) in loop \( C_1 \) causes a certain flux \( \Phi_{21} \) through loop \( C_2 \).
units, with $E$ in volts and $I$ in amperes, the unit for $M_{21}$ is \( \text{volt} \cdot \text{amp}^{-1} \cdot \text{s} \), or ohm \cdot s. This unit is called the \textit{henry};\(^2\)

\[
1 \text{ henry} = 1 \frac{\text{volt} \cdot \text{second}}{\text{amp}} = 1 \text{ ohm} \cdot \text{second}. \tag{7.39}
\]

That is, the mutual inductance $M_{21}$ is one henry if a current $I_1$ changing at the rate of 1 ampere/second induces an electromotive force of 1 volt in circuit $C_2$. In Gaussian units, with $E$ in statvolts and $I$ in esu/second, the unit for $M_{21}$ is statvolt \cdot (esu/second)$^{-1} \cdot \text{second}. Since 1 statvolt equals 1 esu/cm, this unit can also be written as second$^2$/cm.

**Example (Concentric rings)** Figure 7.20 shows two coplanar, concentric rings: a small ring $C_2$ and a much larger ring $C_1$. Assuming $R_2 \ll R_1$, what is the mutual inductance $M_{21}$?

**Solution** At the center of $C_1$, with $I_1$ flowing, the field $B_1$ is given by Eq. (6.54) as

\[
B_1 = \frac{\mu_0 I_1}{2R_1}. \tag{7.40}
\]

Since we are assuming $R_2 \ll R_1$, we can neglect the variation of $B_1$ over the interior of the small ring. The flux through the small ring is then

\[
\Phi_{21} = (\pi R_2^2) \frac{\mu_0 I_1}{2R_1} = \frac{\mu_0 \pi R_2^2}{2R_1}. \tag{7.41}
\]

The mutual inductance $M_{21}$ in Eq. (7.37) is therefore

\[
M_{21} = \frac{\Phi_{21}}{I_1} = \frac{\mu_0 \pi R_2^2}{2R_1}, \tag{7.42}
\]

and the electromotive force induced in $C_2$ is

\[
\mathcal{E}_{21} = -M_{21} \frac{dI_1}{dt} = -\frac{\mu_0 \pi R_2^2}{2R_1} \frac{dI_1}{dt}. \tag{7.43}
\]

Since $\mu_0 = 4\pi \cdot 10^{-7} \text{ kg m/C}^2$, we can write $M_{21}$ alternatively as

\[
M_{21} = \frac{(2\pi^2 \cdot 10^{-7} \text{ kg m/C}^2) R_2^2}{R_1}. \tag{7.44}
\]

The numerical value of this expression gives $M_{21}$ in henrys. In Gaussian units, you can show that the relation corresponding to Eq. (7.43) is

\[
\mathcal{E}_{21} = -\frac{1}{c} \frac{2\pi^2 R_2^2}{cR_1} \frac{dI_1}{dt}, \tag{7.45}
\]

\(^2\) The unit is named after Joseph Henry (1797–1878), the foremost American physicist of his time. Electromagnetic induction was discovered independently by Henry, practically at the same time as Faraday conducted his experiments. Henry was the first to recognize the phenomenon of self-induction. He developed the electromagnet and the prototype of the electric motor, invented the electric relay, and all but invented telegraphy.
with $\mathcal{E}_{21}$ in statvolts, the $R$’s in cm, and $I_1$ in esu/second. $M_{21}$ is the coefficient of the $dI_1/dt$ term, namely $2\pi^2 R_2^2/c^2 R_1$ (in second$^2$/cm). Appendix C states, and derives, the conversion factor from henry to second$^2$/cm.

Incidentally, the minus sign we have been carrying along doesn’t tell us much at this stage. If you want to be sure which way the electromotive force will tend to drive current in $C_2$, Lenz’s law is your most reliable guide.

If the circuit $C_1$ consisted of $N_1$ turns of wire instead of a single ring, the field $B_1$ at the center would be $N_1$ times as strong, for a given current $I_1$. Also, if the small loop $C_2$ consisted of $N_2$ turns, all of the same radius $R_2$, the electromotive force in each turn would add to that in the next, making the total electromotive force in that circuit $N_2$ times that of a single turn. Thus for multiple turns in each coil the mutual inductance will be given by

$$M_{21} = \frac{\mu_0 \pi N_1 N_2 R_2^2}{2 R_1}. \quad (7.46)$$

This assumes that the turns in each coil are neatly bundled together, the cross section of the bundle being small compared with the coil radius. However, the mutual inductance $M_{21}$ has a well-defined meaning for two circuits of any shape or distribution. As we wrote in Eq. (7.38), $M_{21}$ is the (negative) ratio of the electromotive force in circuit 2, caused by changing current in circuit 1, to the rate of change of current $I_1$. That is,

$$M_{21} = -\frac{\mathcal{E}_{21}}{dI_1/dt}. \quad (7.47)$$

### 7.7 A reciprocity theorem

In considering the circuits $C_1$ and $C_2$ in the preceding example, we might have inquired about the electromotive force induced in circuit $C_1$ by a changing current in circuit $C_2$. That would involve another coefficient of mutual inductance, $M_{12}$, given by (ignoring the sign)

$$M_{12} = \frac{\mathcal{E}_{12}}{dI_2/dt}. \quad (7.48)$$

$M_{12}$ is related to $M_{21}$ by the following remarkable theorem.

**Theorem 7.2** For any two circuits,

$$M_{12} = M_{21} \quad (7.49)$$

This theorem is not a matter of geometrical symmetry. Even the simple example in Fig. 7.20 is not symmetrical with respect to the two
circuits. Note that \( R_1 \) and \( R_2 \) enter in different ways into the expression for \( M_{21} \); Eq. (7.49) asserts that, for these two dissimilar circuits, if

\[
M_{21} = \frac{\pi \mu_0 N_1 N_2 R_2^2}{2R_1}, \quad \text{then} \quad M_{12} = \frac{\pi \mu_0 N_1 N_2 R_2^2}{2R_1}
\]

(7.50)

also – and not what we would get by switching 1’s and 2’s everywhere!

**Proof** In view of the definition of mutual inductance in Eq. (7.37), our goal is to show that \( \Phi_{12}/I_2 = \Phi_{21}/I_1 \), where \( \Phi_{12} \) is the flux through some circuit \( C_1 \) due to a current \( I_2 \) in another circuit \( C_2 \), and \( \Phi_{21} \) is the flux through \( C_2 \) due to a current \( I_1 \) in \( C_1 \). We will use the vector potential. Stokes’ theorem tells us that

\[
\int_C A \cdot ds = \int_S (\text{curl } A) \cdot da. \quad (7.51)
\]

In particular, if \( A \) is the vector potential of a magnetic field \( B \), in other words, if \( B = \text{curl } A \), then we have

\[
\int_C A \cdot ds = \int_S B \cdot da = \Phi_s
\]

(7.52)

*That is, the line integral of the vector potential around a loop is equal to the flux of \( B \) through the loop.*

Now, the vector potential is related to its current source as follows, according to Eq. (6.46):

\[
A_{21} = \frac{\mu_0 I_1}{4\pi} \int_{C_1} \frac{ds_1}{r_{21}}, \quad (7.53)
\]

where \( A_{21} \) is the vector potential, at some point \((x_2, y_2, z_2)\), of the magnetic field caused by current \( I_1 \) flowing in circuit \( C_1 \); \( ds_1 \) is an element of the loop \( C_1 \); and \( r_{21} \) is the magnitude of the distance from that element to the point \((x_2, y_2, z_2)\).

**Figure 7.21** shows the two loops \( C_1 \) and \( C_2 \), with current \( I_1 \) flowing in \( C_1 \). Let \((x_2, y_2, z_2)\) be a point on the loop \( C_2 \). Then Eqs. (7.52) and (7.53) give the flux through \( C_2 \) due to current \( I_1 \) in \( C_1 \) as

\[
\Phi_{21} = \int_{C_2} ds_2 \cdot A_{21} = \int_{C_2} ds_2 \cdot \frac{\mu_0 I_1}{4\pi} \int_{C_1} \frac{ds_1}{r_{21}} = \frac{\mu_0 I_1}{4\pi} \int_{C_2} \int_{C_1} \frac{ds_2 \cdot ds_1}{r_{21}}. \quad (7.54)
\]

Similarly, the flux through \( C_1 \) due to current \( I_2 \) flowing in \( C_2 \) is given by the same expression with the labels 1 and 2 reversed:

\[
\Phi_{12} = \frac{\mu_0 I_2}{4\pi} \int_{C_1} \int_{C_2} \frac{ds_1 \cdot ds_2}{r_{12}}. \quad (7.55)
\]
Electromagnetic induction

Now \( r_{12} = r_{21} \), for these are just distance magnitudes, not vectors. The meaning of each of the integrals above is as follows: take the scalar product of a pair of line elements, one on each loop, divide by the distance between them, and sum over all pairs. The only difference between Eqs. (7.54) and (7.55) is the order in which this operation is carried out, and that cannot affect the final sum. Hence \( \Phi_{21}/I_1 = \Phi_{12}/I_2 \), as desired. Thanks to this theorem, we need make no distinction between \( M_{12} \) and \( M_{21} \). We may speak, henceforth, of the mutual inductance \( M \) of any two circuits.

Theorems of this sort are often called “reciprocity” theorems. There are some other reciprocity theorems on electric circuits not unrelated to this one. This may remind you of the relation \( C_{jk} = C_{kj} \) mentioned in Section 3.6 and treated in Exercise 3.64. (In the spirit of that exercise, see Problem 7.10 for a second proof of the above \( M_{12} = M_{21} \) theorem.) A reciprocity relation usually expresses some general symmetry law that is not apparent in the superficial structure of the system.

### 7.8 Self-inductance

When the current \( I_1 \) is changing, there is a change in the flux through circuit \( C_1 \) itself, and consequently an electromotive force is induced. Call this \( \mathcal{E}_{11} \). The induction law holds, whatever the source of the flux:

\[
\mathcal{E}_{11} = -\frac{d\Phi_{11}}{dt},
\]

(7.56)

where \( \Phi_{11} \) is the flux through circuit 1 of the field \( B_1 \) due to the current \( I_1 \) in circuit 1. The minus sign expresses the fact that the electromotive force is always directed so as to oppose the change in current – Lenz’s law, again. Since \( \Phi_{11} \) will be proportional to \( I_1 \) we can write

\[
\frac{\Phi_{11}}{I_1} = \text{constant} \equiv L_1.
\]

(7.57)

Equation (7.56) then becomes

\[
\mathcal{E}_{11} = -L_1 \frac{dI_1}{dt}.
\]

(7.58)

The constant \( L_1 \) is called the self-inductance of the circuit. We usually drop the subscript “1.”

**Example (Rectangular toroidal coil)** As an example of a circuit for which \( L \) can be calculated, consider the rectangular toroidal coil of Exercise 6.61, shown here again in Fig. 7.22. You found (if you worked that exercise) that a current \( I \) flowing in the coil of \( N \) turns produces a field, the strength of which, at a radial distance \( r \) from the axis of the coil, is given by \( B = \mu_0 NI/2\pi r \). The total flux
through one turn of the coil is the integral of this field over the cross section of the coil:

$$\Phi(\text{one turn}) = h \int_a^b \frac{\mu_0 NI}{2\pi r} \, dr = \frac{\mu_0 NIh}{2\pi} \ln \left( \frac{b}{a} \right). \quad (7.59)$$

The flux threading the circuit of \(N\) turns is \(N\) times as great:

$$\Phi = \frac{\mu_0 N^2 h}{2\pi} \ln \left( \frac{b}{a} \right). \quad (7.60)$$

Hence the induced electromotive force \(E\) is

$$E = -\frac{d\Phi}{dt} = -\frac{\mu_0 N^2 h}{2\pi} \ln \left( \frac{b}{a} \right) \frac{dI}{dt}. \quad (7.61)$$

Thus the self-inductance of this coil is given by

$$L = \frac{\mu_0 N^2 h}{2\pi} \ln \left( \frac{b}{a} \right). \quad (7.62)$$

Since \(\mu_0 = 4\pi \cdot 10^{-7} \text{ kg m/C}^2\), we can rewrite this in a form similar to Eq. (7.44):

$$L = (2 \cdot 10^{-7} \text{ kg m/C}^2) N^2 h \ln \left( \frac{b}{a} \right). \quad (7.63)$$

The numerical value of this expression gives \(L\) in henrys. In Gaussian units, you can show that the self-inductance is

$$L = \frac{2N^2 h}{c^2} \ln \left( \frac{b}{a} \right). \quad (7.64)$$

You may think that one of the rings we considered earlier would have made a simpler example to illustrate the calculation of self-inductance. However, if we try to calculate the inductance of a simple circular loop of wire, we encounter a puzzling difficulty. It seems a good idea to simplify the problem by assuming that the wire has zero diameter. But we soon discover that, if finite current flows in a filament of zero diameter, the flux threading a loop made of such a filament is infinite! The reason is that the field \(B\), in the neighborhood of a filamentary current, varies as \(1/r\), where \(r\) is the distance from the filament, and the integral of \(B \times \text{(area)}\) diverges as \(\int (dr/r)\) when we extend it down to \(r = 0\). To avoid this we may let the radius of the wire be finite, not zero, which is more realistic anyway. This may make the calculation a bit more complicated, in a given case, but that won’t worry us. The real difficulty is that different parts of the wire (at different distances from the center of the loop) now appear as different circuits, linked by different amounts of flux. We are no longer sure what we mean by the flux through the circuit. In fact, because the electromotive force is different in the different filamentary loops into which the circuit can be divided, some redistribution of current density must occur when rapidly changing currents flow in the ring. Hence the inductance of the circuit may depend somewhat on the rapidity of change of \(I\), and thus not be strictly a constant as Eq. (7.58) would imply.
We avoided this embarrassment in the toroidal coil example by ignoring the field in the immediate vicinity of the individual turns of the winding. Most of the flux does not pass through the wires themselves, and whenever that is the case the effect we have just been worrying about will be unimportant.

**7.9 Circuit containing self-inductance**

Suppose we connect a battery, providing electromotive force $E_0$, to a coil, or inductor, with self-inductance $L$, as in Fig. 7.23(a). The coil itself, the connecting wires, and even the battery will have some resistance. We don’t care how this is distributed around the circuit. It can all be lumped together in one resistance $R$, indicated on the circuit diagram of Fig. 7.23(b) by a resistor symbol with this value. Also, the rest of the circuit, especially the connecting wires, contribute a bit to the self-inductance of the whole circuit; we assume that this is included in $L$. In other words, Fig. 7.23(b) represents an idealization of the physical circuit. The inductor $L$, symbolized by $\mathbf{L}$, has no resistance; the resistor $R$ has no inductance. It is this idealized circuit that we shall now analyze.

If the current $I$ in the circuit is changing at the rate $dI/dt$, an electromotive force $LdI/dt$ will be induced, in a direction to oppose the change. Also, there is the constant electromotive force $E_0$ of the battery. If we define the positive current direction as the one in which the battery tends to drive current around the circuit, then the net electromotive force at any instant is $E_0 - LdI/dt$. This drives the current $I$ through the resistor $R$. That is,

$$E_0 - L\frac{dI}{dt} = RI. \tag{7.65}$$

We can also describe the situation in this way: the voltage difference between points $A$ and $B$ in Fig. 7.23(b), which we call the voltage across the inductor, is $LdI/dt$, with the upper end of the inductor positive if $I$ in the direction shown is increasing. The voltage difference between $B$ and $C$, the voltage across the resistor, is $RI$, with the upper end of the resistor positive. Hence the sum of the voltage across the inductor and the voltage across the resistor is $LdI/dt + RI$. This is the same as the potential difference between the battery terminals, which is $E_0$ (our idealized battery has no internal resistance). Thus we have

$$E_0 = L\frac{dI}{dt} + RI, \tag{7.66}$$

which is merely a restatement of Eq. (7.65).

Before we look at the mathematical solution of Eq. (7.65), let’s predict what ought to happen in this circuit if the switch is closed at $t = 0$. Before the switch is closed, $I = 0$, necessarily. A long time after the switch has been closed, some steady state will have been attained, with
current practically constant at some value \( I_0 \). Then and thereafter, 
\[ \frac{dI}{dt} \approx 0, \] 
and Eq. (7.65) reduces to 
\[ \mathcal{E}_0 = RI_0. \] 
(7.67)

The transition from zero current to the steady-state current \( I_0 \) cannot occur abruptly at \( t = 0 \), for then \( dI/dt \) would be infinite. In fact, just after \( t = 0 \), the current \( I \) will be so small that the RI term in Eq. (7.65) can be ignored, giving

\[ \frac{dI}{dt} = \frac{\mathcal{E}_0}{L}. \] 
(7.68)

The inductance \( L \) limits the rate of rise of the current.

What we now know is summarized in Fig. 7.24(a). It only remains to find how the whole change takes place. Equation (7.65) is a differential equation very much like Eq. (4.39) in Chapter 4. The constant \( \mathcal{E}_0 \) term complicates things slightly, but the equation is still straightforward to solve. In Problem 7.14 you can show that the solution to Eq. (7.65) that satisfies our initial condition, \( I = 0 \) at \( t = 0 \), is

\[ I(t) = \frac{\mathcal{E}_0}{R} \left(1 - e^{-\left(R/L\right)t}\right). \] 
(7.69)

The graph in Fig. 7.24(b) shows the current approaching its asymptotic value \( I_0 \) exponentially. The “time constant” of this circuit is the quantity \( L/R \). If \( L \) is measured in henrys and \( R \) in ohms, this comes out in seconds, since henrys = volt · amp\(^{-1}\) · second, and ohms = volt · amp\(^{-1}\).

What happens if we open the switch after the current \( I_0 \) has been established, thus forcing the current to drop abruptly to zero? That would make the term \( L \frac{dI}{dt} \) negatively infinite! The catastrophe can be more than mathematical. People have been killed opening switches in highly inductive circuits. What happens generally is that a very high induced voltage causes a spark or arc across the open switch contacts, so that the current continues after all. Let us instead remove the battery from the circuit by closing a conducting path across the LR combination, as in Fig. 7.25(a), at the same time disconnecting the battery. We now have a circuit described by the equation

\[ 0 = L \frac{dI}{dt} + RI, \] 
(7.70)

with the initial condition \( I = I_0 \) at \( t = t_1 \), where \( t_1 \) is the instant at which the short circuit was closed. The solution is the simple exponential decay function

\[ I(t) = I_0 e^{-\left(R/L\right)(t-t_1)}, \] 
(7.71)

with the same characteristic time \( L/R \) as before.
7.10 Energy stored in the magnetic field

During the decay of the current described by Eq. (7.71) and Fig. 7.25(b), energy is dissipated in the resistor $R$. Since the energy $dU$ dissipated in any short interval $dt$ is $RI^2 dt$, the total energy dissipated after the closing of the switch at time $t_1$ is given by

$$U = \int_{t_1}^{\infty} RI^2 dt = \int_{t_1}^{\infty} R I_0^2 e^{-(2R/L)(t-t_1)} dt$$

$$= -RI_0^2 \left( \frac{L}{2R} \right) e^{-(2R/L)(t-t_1)} \bigg|_{t_1}^{\infty} = \frac{1}{2} LI_0^2. \quad (7.72)$$

The source of this energy was the inductor with its magnetic field. Indeed, exactly that amount of work had been done by the battery to build up the current in the first place – over and above the energy dissipated in the resistor between $t = 0$ and $t = t_1$, which was also provided by the battery. To see that this is a general relation, note that, if we have an increasing current in an inductor, work must be done to drive the current $I$ against the induced electromotive force $L dI/dt$. Since the electromotive force is defined to be the work done per unit charge, and since a charge $Idt$ moves through the inductor in time $dt$, the work done in time $dt$ is

$$dW = L \frac{dI}{dt} (Idt) = LI dI = \frac{1}{2} L d(I^2). \quad (7.73)$$

Therefore, we may assign a total energy

$$U = \frac{1}{2} LI^2 \quad (7.74)$$

to an inductor carrying current $I$. With the eventual decay of this current, that amount of energy will appear somewhere else.

It is natural to regard this as energy stored in the magnetic field of the inductor, just as we have described the energy of a charged capacitor as stored in its electric field. The energy of a capacitor charged to potential difference $V$ is $(1/2)CV^2$ and is accounted for by assigning to an element of volume $dv$, where the electric field strength is $E$, an amount of energy $(\epsilon_0/2)E^2 dv$. It is pleasant, but hardly surprising, to find that a similar relation holds for the energy stored in an inductor. That is, we can ascribe to the magnetic field an energy density $(1/2\mu_0)B^2$, and summing the energy of the whole field will give the energy $(1/2)LI^2$.

**Example (Rectangular toroidal coil)** To show how the energy density $B^2/2\mu_0$ works out in one case, we can go back to the toroidal coil whose inductance $L$ we calculated in Section 7.8. We found in Eq. (7.62) that

$$L = \frac{\mu_0 N^2 h}{2\pi} \ln \left( \frac{b}{a} \right). \quad (7.75)$$
The magnetic field strength $B$, with current $I$ flowing, was given by

$$B = \frac{\mu_0 NI}{2\pi r}.$$ (7.76)

To calculate the volume integral of $B^2/2\mu_0$ we can use a volume element consisting of the cylindrical shell sketched in Fig. 7.26, with volume $2\pi rh \, dr$. As this shell expands from $r = a$ to $r = b$, it sweeps through all the space that contains magnetic field. (The field $B$ is zero everywhere outside the torus, remember.) So,

$$\frac{1}{2\mu_0} \int B^2 \, dv = \frac{1}{2\mu_0} \int_a^b \left( \frac{\mu_0 NI}{2\pi r} \right)^2 2\pi rh \, dr = \frac{\mu_0 N^2 h l^2}{4\pi} \ln \left( \frac{b}{a} \right).$$ (7.77)

Comparing this result with Eq. (7.75), we see that, indeed,

$$\frac{1}{2\mu_0} \int B^2 \, dv = \frac{1}{2} LI^2.$$ (7.78)

The task of Problem 7.18 is to show that this result holds for an arbitrary circuit with inductance $L$.

The more general statement, the counterpart of our statement for the electric field in Eq. (1.53), is that the energy $U$ to be associated with any magnetic field $B(x, y, z)$ is given by

$$U = \frac{1}{2\mu_0} \int_{\text{entire field}} B^2 \, dv.$$ (7.79)

With $B$ in tesla and $v$ in m$^3$, the energy $U$ will be given in joules, as you can check. In Eq. (7.74), with $L$ in henrys and $I$ in amperes, $U$ will also be given in joules. The Gaussian equivalent of Eq. (7.79) for $U$ in ergs, $B$ in gauss, and $v$ in cm$^3$ is

$$U = \frac{1}{8\pi} \int_{\text{entire field}} B^2 \, dv.$$ (7.80)

The Gaussian equivalent of Eq. (7.74) remains $U = LI^2/2$, because the reasoning leading up to that equation is unchanged.

### 7.11 Applications

An electrodynamic tether is a long (perhaps 20 km) straight conducting wire that has one end connected to a satellite. The other end hangs down toward (or up away from) the earth. As the satellite and tether orbit the earth, they pass through the earth’s magnetic field. Just as with the moving rod in Section 7.2, an emf is generated along the wire. If this were the whole story, charge would simply pile up on the ends. But the satellite is moving through the ionosphere, which contains enough ions to yield a return path for the charge. A complete circuit is therefore formed, so
the emf can be used to provide power to the satellite. However, the current in the wire will experience a Lorentz force, and you can show that the direction is opposite to the satellite’s motion. On the other hand, if a power source on the satellite drives current in the opposite direction along the tether, the Lorentz force will be in the same direction as the satellite’s motion. So the tether can serve as a (gentle) propulsion device.

If you drop a wire hoop into a region containing a horizontal magnetic field, the changing flux will induce a current in the hoop. From Lenz’s law and the right-hand rule, the resulting Lorentz force on the current is upward, independent of whether the hoop is entering or leaving the region of the magnetic field. So the direction of the force is always opposite to that of the velocity. If you drop a solid metal sheet into the region, loops (or eddies) of current will develop in the sheet, and the same braking effect, known as eddy-current braking, will occur. This braking effect has many applications, from coin vending machines to trains to amusement park rides. The loss in kinetic energy shows up as resistive heating. Eddy currents are also used in metal detectors, both of the airport security type and the hunting-for-buried-treasure type. The metal detector sends out a changing magnetic field, which induces eddy currents in any metal present. These currents produce their own changing magnetic field, which is then detected by the metal detector.

An electric guitar generates its sound via magnetic induction. The strings are made of a material that is easily magnetized, and they vibrate back and forth above pickups. A pickup is a coil of wire wrapped around a permanent magnet. This magnet causes the string to become magnetized, and the string’s magnetic field then produces a flux through the coil. Because the string is vibrating, this flux is changing, so an emf is induced in the coil. This sends an oscillating current (with the same frequency as the string’s vibration) to an amplifier, which then amplifies the sound. Without the amplifier, the sound is barely audible. Electric guitars with nylon (or otherwise nonmagnetic) strings won’t work!

A shake flashlight is a nice application of Faraday’s law. As you shake the flashlight, a permanent magnet passes back and forth through a coil. The induced emf in the coil produces a current that deposits charge on a capacitor, where it can be stored. A bridge rectifier (consisting of a certain configuration of four diodes) changes the alternating-current emf from your shaking motion to a direct-current emf, so that positive charge always flows toward the positive side of the capacitor, independent of which way the magnet is moving through the coil. When you flip the light switch, you allow current to flow from the capacitor through the light bulb.

An electric generator is based on the circuit in Fig. 7.13. An external torque causes the loop of wire to rotate, and the changing flux through the loop induces an oscillating (that is, alternating) emf. In practice, however, in most generators the rotating part (which is effectively the turbine) contains a permanent magnet that produces a changing flux through coils.
of wire arranged around the perimeter. In any case, it is Faraday’s law at work. The force causing the turbine to rotate can come from various sources: water pressure from a dam, air pressure on a windmill, steam pressure from a coal plant or nuclear reactor, etc.

The alternator in a car engine is simply a small electric generator. In addition to turning the wheels, the engine also turns a small magnet inside the alternator, producing an alternating current. However, the car’s battery requires direct current, so a rectifier converts the ac to dc. On an even smaller scale, the magnet in one type of bicycle-light generator is made to rotate due to the friction force from the tire. The power generated is usually only a few watts – a small fraction of the total power output from the rider. But don’t think that you could make much money by harnessing all of your power on a stationary bike. If a professional cyclist sold the power he could generate during one hour of hard pedaling, he would earn about 5 cents.

Hybrid cars, which are powered by both gasoline and a battery, use regenerative braking to capture the kinetic energy of the car when the brakes are applied. The motor acts as a generator. More precisely, the friction force between the tires and the ground provides a torque on the gears in the motor, which in turn provide a torque on the magnet inside the generator.

A microphone is basically the opposite of a speaker (see Section 6.10). It converts sound waves into an electrical signal. A common type, called a dynamic microphone, makes use of electromagnetic induction. A coil of wire is attached to a diaphragm and surrounds one pole of a permanent magnet, in the same manner as in a speaker. When a sound wave causes the diaphragm to vibrate, the coil likewise vibrates. Its motion through the field of the magnet causes changing flux through the coil. An emf, and hence current, are therefore induced in the coil. This current signal is sent to a speaker, which reverses the process, turning the electrical signal back into sound waves. Alternatively, the signal is sent to a device that stores the information; see the discussion of cassette tapes and hard disks in Section 11.12.

A ground-fault circuit interrupter (GFCI) helps prevent (or at least mitigate) electric shocks. Under normal conditions, the current coming out of the “hot” slot in a wall socket (the short slot) equals the current flowing into the neutral slot (the tall slot). Now let’s say you are receiving an electric shock. In a common type of shock, some of the current is taking an alternate route to ground – through you instead of through the neutral wire. The GFCI monitors the difference between the currents in the hot and neutral wires, and if it detects a difference of more than 5 or 10 mA, it trips the circuit (quickly, in about 30 ms). This monitoring is accomplished by positioning a toroidal coil around the two wires to and from the slots. The currents in these wires travel in opposite directions, so when the currents agree, there is zero net current in the pair. But if there is a mismatch in the (oscillating) currents, then the nonzero net
current will produce a changing magnetic field circling around the wires. So there will be a changing magnetic field in the toroidal coil, which will induce a detectable current in the coil. A signal is then sent to a mechanism that trips the circuit. Having survived the shock, you can reset the GFCI by pressing the reset button on the outlet. As with a circuit breaker (see Section 6.10), nothing needs to be replaced after the circuit is tripped. In contrast, a fuse (see Section 4.12) needs to be replaced after it burns out. However, the purpose of a GFCI is different. A GFCI protects people by preventing tiny currents from traveling through them (even 50 mA can disrupt the functioning of a heart), while a fuse or a circuit breaker protects buildings by preventing large currents (on the order of 20 A), which can generate heat and cause fires.

A transformer changes the voltage in a circuit. Imagine two solenoids, A and B, both wound around the same cylinder. Let B (the secondary winding) have ten times the number of turns as A (the primary winding). If a sinusoidal voltage source is connected to A, it will cause a changing flux through A, and hence also through B, with the latter flux being ten times the former. The induced emf in B will therefore be ten times the emf in A. By adjusting the ratio of the number of turns, the voltage can be stepped up or stepped down by any factor. In practice, the solenoids in a transformer aren’t actually right on top of each other, but instead wrapped around different parts of an iron core, which funnels the magnetic field lines along the core from one solenoid to the other. The ease with which ac voltages can be stepped up or down is the main reason why the electric grid uses ac. It is necessary to step up the voltage for long-distance transmission (and hence step down the current, for a given value of the power $P = IV$), because otherwise there would be prohibitively large energy losses due to the $I^2R$ resistance heating in the wires.

An ignition system coil in your car converts the 12 volts from the battery into the 30,000 or so volts needed to cause the arcing (the spark) across the spark plugs. Like a transformer, the ignition coil has primary and secondary windings, but the mechanism is slightly different. Instead of producing an oscillating voltage, it produces a one-time surge in voltage, whose original source was the 12 volt dc battery. It does this in two steps. First, the battery produces a steady current through a circuit containing the primary winding. A switch is then opened, and the current drops rapidly to zero (the rate is controlled by inserting a capacitor in the circuit, in parallel with the switch). This changing current creates a large back emf in the primary coil, say 300 volts. Second, if the secondary winding has, say, 100 times as many turns as the primary, then the transformer reasoning in the preceding paragraph leads to 30,000 volts in the secondary coil. This is enough to cause arcing in the spark plug. This two-step process means that we don’t need to have $30,000/12 \approx 3000$ as many turns in the secondary coil!

A boost converter (or step-up converter) increases the voltage in a dc circuit. It is used, for example, to power a 3 volt LED lamp with a
1.5 volt battery. The main idea behind the converter is the fact that inductors resist sudden changes in current. Consider a circuit where current flows from a battery through an inductor. If a switch is opened downstream from the inductor, and if an alternative path is available through a capacitor, then current will still flow for a brief time through the inductor onto the capacitor. Charge will therefore build up on the capacitor. This process is repeated at a high frequency, perhaps 50 kHz. Even if the back voltage from the capacitor is higher than the forward voltage from the battery, a positive current will still flow briefly onto the capacitor each time the switch is opened. (Backward current can be prevented with a diode.) The capacitor then serves as a higher-voltage effective battery for powering the LED.

The magnetic field of the earth cannot be caused by a permanent magnet, because the interior temperature is far too hot to allow the iron core to exist in a state of permanent magnetization. Instead, the field is caused by the dynamo effect (see Problem 7.19 and Exercise 7.47). A source of energy is needed to drive the dynamo, otherwise the field would decay on a time scale of 20,000 years or so. This source isn’t completely understood; possibilities include tidal forces, gravitational setting, radioactivity, and the buoyancy of lighter elements. The dynamo mechanism requires a fluid region inside the earth (this region is the outer core) and also a means of charge separation (perhaps friction between layers) so that currents can exist. It also requires that the earth be rotating, so that the Coriolis force can act on the fluid. Computer models indicate that the motion of the fluid is extremely complicated, and also that the reversal of the field (which happens every 200,000 years, on average) is likewise complicated. The poles don’t simply rotate into each other. Rather, all sorts of secondary poles appear on the surface of the earth during the process, which probably takes a few thousand years. Who knows where all the famous explorers and their compasses would have ended up if a reversal had been taking place during the last thousand years! In recent years, the magnetic north pole has been moving at the brisk rate of about 50 km per year. This speed isn’t terribly unusual, though, so it doesn’t necessarily imply that a reversal is imminent.

**CHAPTER SUMMARY**

- Faraday discovered that a current in one circuit can be induced by a changing current in another circuit.

- If a loop moves through a magnetic field, the induced emf equals \( \mathcal{E} = vw(B_1 - B_2) \), where \( w \) is the length of the transverse sides, and the \( B \)'s are the fields at these sides. This emf can be viewed as a consequence of the Lorentz force acting on the charges in the transverse sides.
Electromagnetic induction

- More generally, the emf can be written in terms of the magnetic flux as

\[ \mathcal{E} = -\frac{d\Phi}{dt}. \]  

(7.81)

This is known as Faraday’s law of induction, and it holds in all cases: the loop can be moving, or the source of the magnetic field can be moving, or the flux can be changed by some other arbitrary means. The sign of the induced emf is determined by Lenz’s law: the induced current flows in the direction that produces a magnetic field that opposes the change in flux. The differential form of Faraday’s law is

\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \]  

(7.82)

This is one of Maxwell’s equations.

- If we have two circuits \( C_1 \) and \( C_2 \), a current \( I_1 \) in one circuit will produce a flux \( \Phi_{21} \) through the other. The mutual inductance \( M_{21} \) is defined by \( M_{21} = \Phi_{21}/I_1 \). It then follows that a changing \( I_1 \) produces an emf in \( C_2 \) equal to \( \mathcal{E}_{21} = -M_{21} \frac{dI_1}{dt} \). The two coefficients of mutual inductance are symmetric: \( M_{12} = M_{21} \).

- The self-inductance \( L \) is defined analogously. A current \( I \) in a circuit will produce a flux \( \Phi \) through the circuit, and the self-inductance is defined by \( L = \Phi/I \). The emf is then \( \mathcal{E} = -L \frac{dI}{dt} \).

- If a circuit contains an inductor, and if a switch is opened (or closed), the current can’t change discontinuously, because that would create an infinite value of \( \mathcal{E} = -L \frac{dI}{dt} \). The current must therefore gradually change. If a switch is closed in an \( RL \) circuit, the current takes the form

\[ I = \frac{\mathcal{E}_0}{R} \left(1 - e^{-(R/L)t}\right). \]  

(7.83)

The quantity \( L/R \) is the time constant of the circuit.

- The energy stored in an inductor equals \( U = LI^2/2 \). It can be shown that this is equivalent to the statement that a magnetic field contains an energy density of \( B^2/2\mu_0 \) (just as an electric field contains an energy density of \( \varepsilon_0 E^2/2 \)).

Problems

7.1  **Current in a bottle**

An ocean current flows at a speed of 2 knots (approximately 1 m/s) in a region where the vertical component of the earth’s magnetic field is 0.35 gauss. The conductivity of seawater in that region is 4 (ohm-m)^{-1}. On the assumption that there is no other horizontal component of \( \mathbf{E} \) than the motional term \( \mathbf{v} \times \mathbf{B} \), find the density \( J \) of the horizontal electric current. If you were to carry a bottle of seawater through the earth’s field at this speed, would such a current be flowing in it?
7.2 What's doing work? ***

In Fig. 7.27 a conducting rod is pulled to the right at speed \( v \) while maintaining contact with two rails. A magnetic field points into the page. From the reasoning in Section 7.3, we know that an induced emf will cause a current to flow in the counterclockwise direction around the loop. Now, the magnetic force \( qu \times B \) is perpendicular to the velocity \( u \) of the moving charges, so it can't do work on them. However, the magnetic force \( f \) in Eq. (7.5) certainly looks like it is doing work. What's going on here? Is the magnetic force doing work or not? If not, then what is? There is definitely something doing work because the wire will heat up.

7.3 Pulling a square frame **

A square wire frame with side length \( \ell \) has total resistance \( R \). It is being pulled with speed \( v \) out of a region where there is a uniform \( B \) field pointing out of the page (the shaded area in Fig. 7.28). Consider the moment when the left corner is a distance \( x \) inside the shaded area.

(a) What force do you need to apply to the square so that it moves with constant speed \( v \)?

(b) Verify that the work you do from \( x = x_0 \) (which you can assume is less than \( \ell/\sqrt{2} \)) down to \( x = 0 \) equals the energy dissipated in the resistor.

7.4 Loops around a solenoid **

We can think of a voltmeter as a device that registers the line integral \( \int E \cdot ds \) along a path \( C \) from the clip at the end of its (+) lead, through the voltmeter, to the clip at the end of its (−) lead. Note that part of \( C \) lies inside the voltmeter itself. Path \( C \) may also be part of a loop that is completed by some external path from the (−) clip to the (+) clip. With that in mind, consider the arrangement in Fig. 7.29. The solenoid is so long that its external magnetic field is negligible. Its cross-sectional area is 20 cm\(^2\), and the field inside is toward the right and increasing at the rate of 100 gauss/s. Two identical voltmeters are connected to points on a loop that encloses the solenoid and contains two 50 ohm resistors, as shown. The voltmeters are capable of reading microvolts and have high internal resistance. What will each voltmeter read? Make sure your answer is consistent, from every point of view, with Eq. (7.26).

7.5 Total charge **

A circular coil of wire, with \( N \) turns of radius \( a \), is located in the field of an electromagnet. The magnetic field is perpendicular to the coil (that is, parallel to the axis of the coil), and its strength has the constant value \( B_0 \) over that area. The coil is connected by a pair of twisted leads to an external resistance. The total resistance of
Electromagnetic induction

Figure 7.29.

this closed circuit, including that of the coil itself, is \( R \). Suppose the electromagnet is turned off, its field dropping more or less rapidly to zero. The induced electromotive force causes current to flow around the circuit. Derive a formula for the total charge \( Q = \int I \, dt \) that passes through the resistor, and explain why it does not depend on the rapidity with which the field drops to zero.

7.6 Growing current in a solenoid **

An infinite solenoid has radius \( R \) and \( n \) turns per unit length. The current grows linearly with time, according to \( I(t) = Ct \). Use the integral form of Faraday’s law to find the electric field at radius \( r \), both inside and outside the solenoid. Then verify that your answers satisfy the differential form of the law.

7.7 Maximum emf for a thin loop ***

A long straight stationary wire is parallel to the \( y \) axis and passes through the point \( z = h \) on the \( z \) axis. A current \( I \) flows in this wire, returning by a remote conductor whose field we may neglect. Lying in the \( xy \) plane is a thin rectangular loop with two of its sides, of length \( \ell \), parallel to the long wire. The length \( b \) of the other two sides is very small. The loop slides with constant speed \( v \) in the \( \hat{x} \) direction. Find the magnitude of the electromotive force induced in the loop at the moment the center of the loop has position \( x \). For what values of \( x \) does this emf have a local maximum or minimum? (Work in the approximation where \( b \ll x \), so that you can approximate the relevant difference in \( B \) fields by a derivative.)
Figure 9.10.
A standing wave produced by reflection at a perfectly conducting sheet.
properties of the simple plane electromagnetic wave, you can analyze a surprisingly wide variety of electromagnetic devices, including interferometers, rectangular hollow wave guides, and strip lines.

9.6 Energy transport by electromagnetic waves

9.6.1 Power density

The energy the earth receives from the sun has traveled through space in the form of electromagnetic waves that satisfy Eq. (9.18). Where is this energy when it is traveling? How is it deposited in matter when it arrives?

In the case of a static electric field, such as the field between the plates of a charged capacitor, we found that the total energy of the system could be calculated by attributing to every volume element $dv$ an amount of energy $(\varepsilon_0 E^2/2) dv$ and adding it all up. Look back at Eq. (1.53). Likewise, the energy invested in the creation of a magnetic field could be calculated by assuming that every volume element $dv$ in the field contains $(B^2/2\mu_0) dv$ units of energy. See Eq. (7.79). The idea that energy actually resides in the field becomes more compelling when we observe sunlight, which has traveled through a vacuum where there are no charges or currents, making something hot.

We can use this idea to calculate the rate at which an electromagnetic wave delivers energy. Consider a traveling plane wave (not a standing wave) of any form, at a particular instant of time. Assign to every infinitesimal volume element $dv$ an amount of energy $(1/2)(\varepsilon_0 E^2 + B^2/\mu_0) dv$, $\mathbf{E}$ and $\mathbf{B}$ being the electric and magnetic fields in that volume element at that instant. Since $1/\mu_0\varepsilon_0 = c^2$, this energy can be written alternatively as $(\varepsilon_0/2)(E^2 + c^2B^2) dv$. Now assume that this energy simply travels with speed $c$ in the direction of propagation. In this way we can find the amount of energy that passes, per unit time, through unit area perpendicular to the direction of propagation. Let us apply this to the sinusoidal wave described by Eqs. (9.22) and (9.23). At the instant $t = 0$, we have $E^2 = E_0^2 \sin^2 y$. Also, $B^2 = (E_0/c)^2 \sin^2 y$, since, as we subsequently found, $B_0$ must equal $\pm E_0/c$. The energy density in this field is therefore

$$\frac{\varepsilon_0}{2} \left( E_0^2 \sin^2 y + c^2 \left( \frac{E_0}{c} \right)^2 \sin^2 y \right) = \varepsilon_0 E_0^2 \sin^2 y.$$ (9.33)
continuous, repetitive wave, whether sinusoidal or not, the rate of energy flow per unit area, which we call the \textit{power density} $S$, is given by

$$S = \epsilon_0 E^2 c$$  \hfill (9.34)$$

Here $E^2$ is the mean square electric field strength, which was $E_0^2/2$ for the sinusoidal wave of amplitude $E_0$. $S$ will be in joules per second per square meter, or equivalently watts per square meter, if $E$ is in volts per meter and $c$ is in meters per second.

In Gaussian units the formula for power density is

$$S = \frac{E^2 c}{4\pi},$$  \hfill (9.35)$$

where $S$ is in ergs per second per square centimeter if $E$ is in statvolts per centimeter and $c$ is in centimeters per second.

If you want to write Eq. (9.34) without reference to $c$, then substituting $c = 1/\sqrt{\mu_0\epsilon_0}$ yields

$$S = \frac{E^2}{\sqrt{\mu_0/\epsilon_0}}$$  \hfill (9.36)$$

This expression for $S$ is based only on the physics that was known in 1861 when Maxwell wrote down his set of equations. That is, it invokes nothing about the nature of light; you can repeat the above derivation by using the expression for $v$ in Eq. (9.26) without introducing the speed of light, $c$. The fact that $1/\sqrt{\mu_0\epsilon_0}$ can indeed be replaced by $c$ was conjectured by Maxwell in 1862, demonstrated experimentally by Hertz in 1888, and explained theoretically by Einstein in 1905 through his special theory of relativity. The last of these routes was the one we took in Chapters 5 and 6, where we showed that $\mu_0 = 1/\epsilon_0 c^2$.

The constant $\sqrt{\mu_0/\epsilon_0}$ in Eq. (9.36) has the dimensions of resistance, and its value is 376.73 ohms. Rounding it off to 377 ohms, we have a convenient and easily remembered formula:

$$S(\text{watts/meter}^2) = \frac{E^2 (\text{volts/meter})^2}{377 \text{ ohms}}$$  \hfill (9.37)$$

The units here reduce to: watts = volt$^2$/ohm, which are the same as in the standard $\text{P} = \text{V}^2/\text{R}$ expression for the power in an ordinary resistor. If you need help in remembering the number 377, it happens to be the number of radians per second in 60 hertz, and also the 14th Fibonacci number.

When the electromagnetic wave encounters an electrical conductor, the electric field causes currents to flow. This generally results in energy being dissipated within the conductor at the expense of the energy in the wave. The total reflection of the incident wave in Fig. 9.10 was a special case in which the conductivity of the reflecting surface was infinite.
If the resistivity of the reflector is not zero, the amplitude of the reflected wave will be less than that of the incident wave. Aluminum, for example, reflects visible light, at normal incidence, with about 92 percent efficiency. That is, 92 percent of the incident energy is reflected, the amplitude of the reflected wave being $\sqrt{0.92}$ or 0.96 times that of the incident wave. The lost 8 percent of the incident energy ends up as heat in the aluminum, where the current driven by the electric field of the wave encounters ohmic resistance. What counts, of course, is the resistivity of aluminum at the frequency of the light wave, in this case about $5 \cdot 10^{14}$ Hz. That may be somewhat different from the dc or low-frequency resistivity of the metal. Still, the reflectivity of most metals for visible light is essentially due to the same highly mobile conduction electrons that make metals good conductors of steady current. It is no accident that good conductors are generally shiny. But why clean copper looks reddish while aluminum looks “silvery” can’t be explained without a detailed theory of each metal’s electronic structure.

Energy can also be absorbed when an electromagnetic wave meets nonconducting matter. Little of the light that strikes a black rubber tire is reflected, although the rubber is an excellent insulator for low-frequency electric fields. Here the dissipation of the electromagnetic energy involves the action of the high-frequency electric field on the electrons in the molecules of the material. In the broadest sense, that applies to the absorption of light in everything around us, including the retina of the eye.

Some insulators transmit electromagnetic waves with very little absorption. The transparency of glass for visible light, with which we are so familiar, is really a remarkable property. In the purest glass fibers used for optical transmission of audio and video signals, a wave travels as much as a hundred kilometers, or more than $10^{11}$ wavelengths, before most of the energy is lost. However transparent a material medium may be, the propagation of an electromagnetic wave within the medium differs in essential ways from propagation through the vacuum. The matter interacts with the electromagnetic field. To take that interaction into account, Eq. (9.18) must be modified in a way that will be explained in Chapter 10.

### 9.6.2 The Poynting vector

With the help of Maxwell’s equations, we can produce a more general version of the power density given in Eq. (9.34). That result was valid only for traveling waves. The present result will be valid for arbitrary electromagnetic fields. Furthermore, it will be valid as a function of time (and space), and not just as a time average. As above, our starting point will be the fact that the energy density of an electromagnetic field, which we label $U$, is given by $\epsilon_0 E^2/2 + B^2/2\mu_0$. Consider the rate of change of $U$. If we write $E^2$ and $B^2$ as $E \cdot E$ and $B \cdot B$, then

$$\frac{\partial U}{\partial t} = \epsilon_0 \frac{\partial E}{\partial t} \cdot E + \frac{1}{\mu_0} \frac{\partial B}{\partial t} \cdot B.$$  
(9.38)
The product rule works for vectors just as it does for regular functions, as you can check by explicitly writing out the Cartesian components. We can rewrite the time derivatives here with the help of the two “induction” Maxwell equations in free space, \( \nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \) and \( \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \). This yields
\[
\frac{\partial U}{\partial t} = \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \cdot \mathbf{E} - \frac{1}{\mu_0} (\nabla \times \mathbf{E}) \cdot \mathbf{B}.
\] (9.39)

The right-hand side of this expression conveniently has the same form as the right-hand side of the vector identity
\[
\nabla \cdot (\mathbf{C} \times \mathbf{D}) = (\nabla \times \mathbf{C}) \cdot \mathbf{D} - (\nabla \times \mathbf{D}) \cdot \mathbf{C}.
\] (9.40)

Hence \( \frac{\partial U}{\partial t} = (1/\mu_0) \nabla \cdot (\mathbf{B} \times \mathbf{E}) \). For reasons that will become clear, let’s switch the order of \( \mathbf{B} \) and \( \mathbf{E} \), which brings in a minus sign. We then have
\[
\frac{\partial U}{\partial t} = -\frac{1}{\mu_0} \nabla \cdot (\mathbf{E} \times \mathbf{B}).
\] (9.41)

If we now define the *Poynting vector* \( \mathbf{S} \) by
\[
\mathbf{S} \equiv \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}
\] (Poynting vector),

then we can write our result as
\[
-\frac{\partial U}{\partial t} = \nabla \cdot \mathbf{S}.
\] (9.43)

This equation should remind you of another one we have encountered. It has exactly the same form as the continuity equation,
\[
-\frac{\partial \rho}{\partial t} = \nabla \cdot \mathbf{J}.
\] (9.44)

Therefore, just as \( \mathbf{J} \) gives the current density (the flow of charge per time per area), we can likewise say that \( \mathbf{S} \) gives the power density (the flow of energy per time per area). Equivalently, Eqs. (9.43) and (9.44) are the statements of conservation of energy and charge, respectively. Energy (or charge) can’t just disappear; if the energy in a given region decreases, it must be the case that energy flowed out of that region, and into another region.

If you don’t trust the analogy with \( \mathbf{J} \), you can work with the integral form of Eq. (9.43). The integral of the energy density \( U \) over a given volume \( V \) is simply the total energy \( U \) contained in that volume. So we have
\[
\frac{dU}{dt} = \frac{d}{dt} \int_V U \, dv = \int_V \frac{\partial U}{\partial t} \, dv = -\int_V \nabla \cdot \mathbf{S} \, dv = -\int_S \mathbf{S} \cdot d\mathbf{a},
\] (9.45)
Maxwell’s equations and EM waves

where we have used the divergence theorem. This shows that the rate of change of the energy in a given volume $V$ equals the negative of the flux of the vector $\mathbf{S}$ outward through the closed surface $S$ that bounds $V$. (Remember that $\mathbf{da}$ is defined to be the outward-pointing normal.) The minus sign in Eq. (9.45) makes sense; a positive outward flux of $\mathbf{S}$ means that $U$ is decreasing. Since Eq. (9.45) holds for an arbitrary closed volume, the natural interpretation of $\mathbf{S}$ is that it gives the rate of energy flow per area through any surface, closed or not.

The Poynting vector $\mathbf{S}$ gives the power density for an arbitrary electromagnetic field, not just for the special case of a traveling wave. For any electromagnetic field, at any given point at any instant in time, the direction of $\mathbf{S}$ gives the direction of the energy flow, and the magnitude of $\mathbf{S}$ gives the energy per time per area flowing through a small frame. The units of $\mathbf{S}$ are joules per second per square meter, or watts per square meter.

In the special case of a traveling wave (sinusoidal or not), we know from the third property listed in Section 9.4 that the velocity points in the direction of $\mathbf{E} \times \mathbf{B}$. This equals the direction of $\mathbf{S}$, as must be the case. We also know that a traveling wave has $\mathbf{B}$ perpendicular to $\mathbf{E}$, with $B = E/c$. The magnitude of $\mathbf{S}$ is therefore $S = E(E/c)/\mu_0$. Using $\mu_0 = 1/\epsilon_0 c^2$, we obtain $S = \epsilon_0 E^2 c$. This is the instantaneous power density. Its average value is simply $\overline{S} = \epsilon_0 E^2 c$, in agreement with Eq. (9.34). (In that equation we were using $\overline{S}$, without the line over it, to denote the average power density.)

Interestingly, there can also be energy flow in a static electromagnetic field. Consider a very long stick with uniform linear charge density $\lambda$, moving with speed $v$ in the longitudinal direction, say, rightward. Close to the stick and not too close to the ends, the stick creates $\mathbf{E}$ and $\mathbf{B}$ fields that are essentially static, with $\mathbf{E}$ pointing radially and $\mathbf{B}$ pointing tangentially. Their cross product is therefore nonzero, so the Poynting vector is nonzero. Hence there is energy flow, and it moves in the same direction as the stick moves (for either sign of $\lambda$), as you can show with the right-hand rule. The energy density at a given point (not too close to the ends) doesn’t change, because energy flows into a given volume from the left at the same rate it flows out to the right. However, near the ends the fields are changing, so there is a net energy flow into or out of a given volume. (Think of a uniform caravan of cars moving along the highway. The density of cars changes only at points near the ends of the caravan.) The rightward flow of energy is consistent with the fact that the whole system is moving to the right.

The Poynting vector (named after John Henry Poynting) falls into a wonderful class of phonetically accurate theorems/results. Others are the Low energy theorem (after F. E. Low) dealing with low-energy photons, and the Schwarzschild radius of a black hole (after Karl Schwarzschild, whose last name means “black shield” in German).
Example (Energy flow into a capacitor) A capacitor has circular plates with radius $R$ and is being charged by a constant current $I$. The electric field $E$ between the plates is increasing, so the energy density is also increasing. This implies that there must be a flow of energy into the capacitor. Calculate the Poynting vector at radius $r$ inside the capacitor (in terms of $r$ and $E$), and verify that its flux equals the rate of change of the energy stored in the region bounded by radius $r$.

Solution If the Poynting vector is to be nonzero, there must be a nonzero magnetic field inside the capacitor. And indeed, because the electric field is changing, there is an induced magnetic field due to the \( \nabla \times \mathbf{B} = \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \) Maxwell equation. If we integrate this equation over the area of a disk with radius $r$ inside the capacitor (see Fig. 9.11) and use Stokes’ theorem on the left-hand side, we obtain

\[
\int \mathbf{B} \cdot d\mathbf{s} = \epsilon_0 \mu_0 \frac{\partial E}{\partial t} \text{(area)} \implies B(2\pi r) = \epsilon_0 \mu_0 \frac{\partial E}{\partial t}(\pi r^2)
\]

\[
\implies B = \frac{\epsilon_0 \mu_0 r}{2} \frac{\partial E}{\partial t}. \quad (9.46)
\]

This magnetic field points tangentially around the circle of radius $r$. Since $E$ is increasing upward, $\mathbf{B}$ is directed counterclockwise when viewed from above, as you can check via the right-hand rule. The Poynting vector $\mathbf{S} = (\mathbf{E} \times \mathbf{B})/\mu_0$ then points radially inward everywhere on the circle of radius $r$. So the direction is correct; energy is flowing into the region bounded by radius $r$.

Let’s now find the magnitude of $\mathbf{S}$. Since $\mathbf{E}$ is perpendicular to $\mathbf{B}$, the magnitude of $\mathbf{S}$ is

\[
S = \frac{EB}{\mu_0} = \frac{E}{\mu_0} \left( \frac{\epsilon_0 \mu_0 r}{2} \frac{\partial E}{\partial t} \right) = \frac{\epsilon_0 r}{2} E \frac{\partial E}{\partial t}. \quad (9.47)
\]

To find the total energy per time (that is, the power) flowing past radius $r$, we must multiply $S$ by the lateral area of the cylinder of radius $r$; that is, we must find the flux of $S$. If the separation between the plates is $h$, the lateral area is $2\pi rh$. The total power flowing into the cylinder of radius $r$ is then

\[
P = \left( \frac{\epsilon_0 r}{2} E \frac{\partial E}{\partial t} \right) 2\pi rh = (\pi r^2 h) \epsilon_0 E \frac{\partial E}{\partial t} = \frac{d}{dt} \left( \text{(volume)} \frac{\epsilon_0 E^2}{2} \right) = \frac{dU}{dt}. \
\]

\[
(9.48)
\]

So the Poynting-vector flux does indeed equal the rate of change of the stored energy. In the special case where $r$ equals the radius of the capacitor, $R$, we obtain the total power flowing into the capacitor. Note that $S$ and $P$ are largest at $r = R$, and zero at $r = 0$, as expected.

Remark: You might be worried that although we found there to be a nonzero magnetic field inside the capacitor, we didn’t take into account the resulting magnetic energy density, $B^2/2\mu_0$. We used only the electric $\epsilon_0 E^2/2$ part of the
density. However, a constant current $I$ implies a constant $d\sigma/dt$ (where $\pm\sigma$ are the charge densities on the plates), which in turn implies a constant $\partial E/\partial t$, which in turn implies a constant $B$, from Eq. (9.46). The magnetic energy density is therefore constant and thus doesn’t affect the $dU/dt$ in Eq. (9.48). We can therefore rightfully ignore it. On the other hand, if $I$ isn’t constant, then things are more complicated. However, for “everyday” rates of change of $I$, it is a very good approximation to say that the magnetic energy density in a capacitor is much smaller than the electric energy density; see Exercise 9.30.

At the end of Section 4.3 we mentioned that the energy flow in a circuit is due to the Poynting vector. We can now say more about this. There are two important parts to the energy flow. The first is the flow that yields the resistance heating. The current in a conducting wire is caused by a longitudinal $E$ field inside the wire; recall $J = \sigma E$. Since the curl of $E$ is zero, this same longitudinal $E$ component must also exist right outside the surface of the wire. As you can show in Exercise 9.28, the Poynting-vector flux through a cylinder right outside the wire exactly accounts for the $IV$ resistance heating.

The second part is the energy flow along the wire. As discussed at the end of Section 4.3, there are surface charges on the wire. These create an electric field perpendicular to the wire, which in turn creates a Poynting vector parallel to the wire, as you can verify. This gives an energy flow along the wire; see Galili and Goihbarg (2005). More generally, the energy flow need not be constrained to lie near the wire if the wire loops around in space. Energy can flow across open space too, from one part of a circuit to another; see Jackson (1996).

If there are other electric fields present in the system, there can be a third part to the energy flow, now away from the wire. See Problem 9.10.

### 9.7 How a wave looks in a different frame

A plane electromagnetic wave is traveling through the vacuum. Let $\mathbf{E}$ and $\mathbf{B}$ be the electric and magnetic fields measured at some place and time in $F$, by an observer in $F$. What field will be measured by an observer in a different frame who happens to be passing that point at that time? Suppose that frame $F'$ is moving with speed $v$ in the $\hat{x}$ direction relative to $F$, with its axes parallel to those of $F$. We can turn to Eq. (6.74) for the transformations of the field components. Let us write them out again:

\[
\begin{align*}
E_x' &= E_x, & E_y' &= \gamma(E_y - vB_z), & E_z' &= \gamma(E_z + vB_y); \\
B_x' &= B_x, & B_y' &= \gamma(B_y + (v/c^2)E_z), & B_z' &= \gamma(B_z - (v/c^2)E_y). 
\end{align*}
\] (9.49)
The key to our problem is the way two particular scalar quantities transform, namely, \( E \cdot B \) and \( E^2 - B^2 \). Let us use Eq. (9.49) to calculate \( E' \cdot B' \) and see how it is related to \( E \cdot B \):

\[
E' \cdot B' = E'_x B'_x + E'_y B'_y + E'_z B'_z \\
= E_x B_x + \gamma^2 \left[ E_y B_y + (v/c^2)E_y B_z - \nu B_y B_z - (v/c)^2 E_z B_z \right] \\
+ \gamma^2 \left[ E_z B_z - (v/c^2)E_z B_x + \nu B_z B_x - (v/c)^2 E_y B_y \right] \\
= E_x B_x + \gamma^2 (1 - \beta^2)(E_y B_y + E_z B_z) = E \cdot B.
\] (9.50)

The scalar product \( E \cdot B \) is not changed in the Lorentz transformation of the fields; it is an invariant. A similar calculation, which will be left to the reader as Exercise 9.32, shows that \( E^2 - c^2 B^2 \) is also unchanged by the Lorentz transformation. We therefore have

\[
E' \cdot B' = E \cdot B \quad \text{and} \quad E'^2 - c^2 B'^2 = E^2 - c^2 B^2
\] (9.51)

The invariance of these two quantities is an important general property of any electromagnetic field, not just the field of an electromagnetic wave with which we are concerned at the moment. For the wave field, its implications are especially simple and direct. We know that the plane wave has \( B \) perpendicular to \( E \), and \( cB = E \). Each of our two invariants, \( E \cdot B \) and \( E^2 - c^2 B^2 \), is therefore zero. And if an invariant is zero in one frame, it must be zero in all frames. We see that any Lorentz transformation of the wave will leave \( E \) and \( cB \) perpendicular and equal in magnitude. A light wave looks like a light wave in any inertial frame of reference. That should not surprise us. It could be said that we have merely come full circle, back to the postulates of relativity, Einstein’s starting point. Indeed, according to Einstein’s own autobiographical account, he had begun 10 years earlier (at age 16!) to wonder what one would observe if one could “catch up” with a light wave. With the transformations in Eq. (9.49), which were given in Einstein’s 1905 paper, the question can be answered. Consider a traveling wave with amplitudes given by

\[
E_y = E_0, \quad E_x = E_z = 0, \quad B_z = E_0/c, \quad B_x = B_y = 0.
\]

This is a wave traveling in the \( \hat{x} \) direction, as we can tell from the fact that \( E \times B \) points in that direction. Using Eq. (9.49) and the relation \( \gamma = 1/\sqrt{1 - \beta^2} \), we find that

\[
E'_y = E_0 \sqrt{\frac{1 - \beta}{1 + \beta}}, \quad B'_z = \frac{E_0}{c} \sqrt{\frac{1 - \beta}{1 + \beta}}.
\] (9.52)

As observed in \( F' \) the amplitude of the wave is reduced. The wave velocity, of course, is \( c \) in \( F' \), as it is in \( F \). The electromagnetic wave has no rest frame. In the limit \( \beta = 1 \), the amplitudes \( E'_y \) and \( B'_z \) observed in \( F' \) are reduced to zero. The wave has vanished!
9.8 Applications

The power density of sunlight when it reaches the earth (or rather, the top of the atmosphere) is about 1360 W/m², on average. You can show that this implies that the total power output of the sun is about $4 \cdot 10^{26}$ W. If the power in one square kilometer of sunlight were converted to electricity at 15 percent efficiency, the result would be 200 megawatts. However, assuming that the sun is shining for only 6 hours per day on average, this would yield an average of 50 megawatts. Effects of atmosphere absorption and latitude would further reduce the result somewhat, but there would still be enough electrical power for a city of 25,000 people.

The cosmic microwave background (CMB) radiation (see Exercise 9.25) was discovered by Penzias and Wilson in 1965. This radiation is left over from the big bang and fills all of space. About 300,000 years after the big bang, the universe became transparent to photons, shortly after the hot plasma of electrons and ions cooled to the point where stable atoms could form. The CMB photons have been traveling freely ever since. The wavelength was shorter back then, but it has been continually expanding along with the expansion of the universe. The radiation consists of a distribution of wavelengths, but the peak is around 2 mm. It looks nearly the same in all directions, but its slight anisotropies yield information about what the early universe looked like.

Comets generally have two kinds of tails. The dust tail consists of dust that is pushed away from the comet by the radiation pressure from the sunlight. (The sunlight carries energy, so it also carries momentum; see Problem 9.11.) The dust drifts relatively slowly away from the comet, so this results in the tail curving and drifting behind the comet. The ion tail consists of ions that are blown away from the comet by the sun’s solar wind (consisting of charged particles). These ions move very quickly away from the comet, so the ion tail always points essentially radially away from the sun, independent of the comet’s location around the sun.

Radio frequency identification (RFID) tags have many uses: anti-theft tags, inventory tracking, tollbooth transponders, chip timing in road races, library books, and so on. Although some RFID tags contain their own power source, most (called “passive RFID”) do not. They are powered by resonant inductive coupling: a small coil and capacitor in the tag constitute an LC circuit with a particular resonant frequency. A “reader” transmits a radio wave with this frequency, and the (changing) magnetic field in this wave induces a current in the tag’s circuit. This powers a small microchip, which then transmits a specific identification message back to the reader.

Cell phones, radios, and many other communication devices make use of electromagnetic waves in the radio frequency part of the spectrum, usually from about 1 MHz to a few GHz. A pure sinusoidal wave at a given frequency contains minimal information, so if we want to transmit
useful information, we must modify the wave in some manner. The two simplest ways of modifying the wave are amplitude modulation (AM) and frequency modulation (FM). In the case of AM, the carrier wave (with a frequency in the 1 MHz range) has its amplitude modulated by the sound wave (with a much smaller frequency in the 1 kHz range) that is being sent. The larger the value of the sound wave at a given instant, the larger the amplitude of the transmitted wave. The sound wave is in some sense the envelope of the transmitted wave. The receiver is able to extract the amplitude information and can then reconstruct the original sound wave; a plot of the amplitude of the transmitted wave as a function of time is effectively a plot of the original sound wave as a function of time.

In the case of FM, the carrier wave (with a frequency of around 100 MHz, as you know from your FM radio dial) has its frequency modulated by the sound wave that is being sent. The larger the value of the sound wave at a given instant, the more the carrier-wave frequency shifts relative to a particular value. The receiver is able to extract the frequency information and can then reconstruct the original sound wave; a plot of the frequency as a function of time is effectively a plot of the original sound wave as a function of time. Note that this frequency is well defined, because even if a time interval is fairly short on the time scale of the sound wave, a very large number of the carrier-wave oscillations still fit into it. One method of extracting the frequency information is called slope detection. In this method, the resonant frequency of the receiver is chosen to be slightly shifted from that of the carrier wave, so that the span of the carrier wave’s frequencies lies on the steep side part of the resonance peak. If the span lies, say, on the left side of the peak, then the response of the receiver’s circuit increases (approximately linearly) as the frequency of the carrier wave increases. So we simply need to measure the amplitude of the current in the circuit, and we will obtain the frequency of the carrier wave (up to some factor). The thing that makes all of this possible is resonance, which enables the receiver to respond to a narrow range of frequencies and ignore all others.

**CHAPTER SUMMARY**

- Because the divergence of the curl of a vector is identically zero, the differential form of Ampère’s law, \( \text{curl} \mathbf{B} = \mu_0 \mathbf{J} \), implies that \( \text{div} \mathbf{J} \) is identically zero. This is inconsistent with the continuity equation, \( \text{div} \mathbf{J} = -\frac{\partial \rho}{\partial t} \) (which follows from conservation of charge), in situations where \( \rho \) changes with time. Therefore, \( \text{curl} \mathbf{B} = \mu_0 \mathbf{J} \) cannot be correct. The correct expression has an extra term, \( \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \), on the right-hand side. With this term, \( \text{div} \mathbf{J} \) correctly equals \( -\frac{\partial \rho}{\partial t} \).
• The quantity \( \epsilon_0 \partial E/\partial t \) is called the \textit{displacement current}. This is the last piece of the puzzle, and we can now write down the complete set of \textit{Maxwell’s equations}:

\[
\begin{align*}
\text{curl } E &= -\frac{\partial B}{\partial t}, \\
\text{curl } B &= \mu_0 \epsilon_0 \frac{\partial E}{\partial t} + \mu_0 \mathbf{J}, \\
\text{div } E &= \frac{\rho}{\epsilon_0}, \\
\text{div } B &= 0. 
\end{align*}
\]

These are, respectively, (1) \textit{Faraday’s law}, (2) \textit{Ampère’s law} with the addition of the displacement current, (3) \textit{Gauss’s law}, and (4) the statement that there are no magnetic monopoles.

• A possible form of a \textit{traveling electromagnetic wave} is

\[
\begin{align*}
E &= \hat{z} E_0 \sin(y - vt) \quad \text{and} \quad B = \hat{x} B_0 \sin(y - vt), 
\end{align*}
\]

where

\[
v = \pm \frac{1}{\sqrt{\mu_0 \epsilon_0}} = \pm c \quad \text{and} \quad E_0 = \pm \frac{B_0}{\sqrt{\mu_0 \epsilon_0}} = \pm c B_0.
\]

In general, we can produce a traveling wave by replacing the \( \sin(y - vt) \) function with \textit{any} function \( f(y - vt) \), provided that (1) \( v = \pm c \), (2) \( E_0 = \pm c B_0 \), and (3) \( \mathbf{E} \) and \( \mathbf{B} \) are perpendicular to each other and also to the direction of propagation.

• A \textit{standing wave} is formed by adding two waves traveling in opposite directions. In a standing wave there are (unlike in a traveling wave) positions where \( \mathbf{E} \) is zero at all times, and times when \( \mathbf{E} \) is zero at all positions. Likewise for \( \mathbf{B} \).

• The \textit{power density} (energy per unit area per unit time) of a sinusoidal electromagnetic wave can be written in various forms:

\[
S = \epsilon_0 E^2 c = \frac{E^2}{\sqrt{\mu_0 / \epsilon_0}} = \frac{E^2 (\text{volts/meter})^2}{377 \text{ ohms}}. 
\]

More generally, the \textit{Poynting vector},

\[
S = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0},
\]

gives the power density of an arbitrary electromagnetic field at every point.

• Using the fact that the \( \mathbf{E} \) and \( \mathbf{B} \) fields transform according to the Lorentz transformations, we can derive two invariants:

\[
E' \cdot B' = E \cdot B \quad \text{and} \quad E'^2 - c^2 B'^2 = E^2 - c^2 B^2. 
\]

These imply that if in one frame \( \mathbf{E} \) and \( \mathbf{B} \) are perpendicular and \( E = c B \), then these two relations are also true in any other frame. That is,
a light wave in one inertial frame looks like a light wave in any other inertial frame.

**Problems**

9.1 *The missing term* **
Due to the contradiction between Eqs. (9.2) and (9.5), we know that there must be an extra term in the $\nabla \times \mathbf{B}$ relation, as we found in Eq. (9.10). Call this term $W$. In the text, we used the Lorentz transformations to motivate a guess for $W$. Find $W$ here by taking the divergence of both sides of $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + W$. Assume that the only facts you are allowed to work with are (1) $\nabla \cdot \mathbf{E} = \rho/\varepsilon_0$, (2) $\nabla \cdot \mathbf{B} = 0$, (3) $\nabla \cdot \mathbf{J} = -\partial \rho/\partial t$, and (4) $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ in the case of steady currents.

9.2 *Spherically symmetric current* *
A spherically symmetric (and constant) current density flows radially inward to a spherical shell, causing the charge on the shell to increase at the constant rate $dQ/dt$. Verify that Maxwell’s equation, $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \partial \mathbf{E}/\partial t$, is satisfied at points outside the shell.

9.3 *A charge and a half-infinite wire* **
A half-infinite wire carries current $I$ from negative infinity to the origin, where it builds up at a point charge with increasing $q$ (so $dq/dt = I$). Consider the circle shown in Fig. 9.12, which has radius $b$ and subtends an angle $2\theta$ with respect to the charge. Calculate the integral $\int \mathbf{B} \cdot ds$ around this circle. Do this in three ways.

(a) Find the $\mathbf{B}$ field at a given point on the circle by using the Biot–Savart law to add up the contributions from the different parts of the wire.

(b) Use the integrated form of Maxwell’s equation (that is, the generalized form of Ampère’s law including the displacement current),

$$\int_C \mathbf{B} \cdot ds = \mu_0 I + \mu_0 \varepsilon_0 \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{a},$$

(9.59)

with $S$ chosen to be a surface that is bounded by the circle and doesn’t intersect the wire, but is otherwise arbitrary. (You can invoke the result from Problem 1.15.)

(c) Use the same strategy as in (b), but now let $S$ intersect the wire.

9.4 *$B$ in a discharging capacitor, via conduction current* **
As mentioned in Exercise 9.15, the magnetic field inside a discharging capacitor can be calculated by summing the contributions from all elements of conduction current. This calculation is
extremely tedious. We can, however, get a handle on the contribution from the plates’ conduction current in a much easier way that doesn’t involve a nasty integral. If we make the usual assumption that the distance $s$ between the plates is small compared with their radius $b$, then any point $P$ inside the capacitor is close enough to the plates so that they look essentially like infinite planes, with a surface current density equal to the density at the nearest point.

(a) Determine the current that crosses a circle of radius $r$ in the capacitor plates, and then use this to find the surface current density. *Hint:* The charge on each plate is essentially uniformly distributed at all times.

(b) Combine the field contributions from the wires and the plates to show that the field at a point $P$ inside the capacitor, a distance $r$ from the axis of symmetry, equals $B = \mu_0 I r / (2\pi b^2)$. (Assume $s \ll r$, so that you can approximate the two wires as a complete infinite wire.)

### 9.5 Maxwell’s equations for a moving charge ***

In part (b) of Problem 6.24 we dealt with approximate expressions for the electric and magnetic fields due to a slowly moving charge, valid in the limit $v \ll c$. In this problem we will use the exact forms. The exact $E$, for any value of $v$, is given in Eq. (5.13) or Eq. (5.15). The Lorentz transformation then gives the exact $B$ as $B = (1/c^2) v \times E$; see Problem 6.24(a). Verify that these exact expressions for $E$ and $B$ satisfy Maxwell’s equations in vacuum. That is:

(a) Show that $\nabla \cdot B = 0$. (We already showed in Problem 5.4 that $\nabla \cdot E = 0$.) The vector identity for $\nabla \cdot (A \times B)$ in Appendix K will come in handy.

(b) Show that $\nabla \times E = -\partial B / \partial t$. (The $\nabla \times B = \partial E / \partial t$ calculation is nearly the same, so you can skip that.) *Note:* Although the calculation is doable if you use the spherical-coordinate expression for $E$ in Eq. (5.15) (don’t forget that both $r$ and $\theta$ vary with time), it’s a bit easier if you use the Cartesian-coordinate expression in Eq. (5.13).

### 9.6 Oscillating field in a solenoid ***

A solenoid with radius $R$ has $n$ turns per unit length. The current varies with time according to $I(t) = I_0 \cos \omega t$. The magnetic field inside the solenoid, $B(t) = \mu_0 n I(t)$, therefore changes with time. In this problem you will need to make use of wisely chosen Faraday/Ampère loops.

---

2 If you want to derive the magnetic field for a fast-moving charge via the Biot–Savart law, you need to incorporate the so-called “retarded time” arising from the finite speed of light. We won’t get into that here, but see Problem 6.28 for a special case.
Problems

(a) Changing $B$ fields cause $E$ fields. Assuming that the $B$ field is given by $B_0(t) = \mu_0 n I_0 \cos \omega t$, find the electric field at radius $r$ inside the solenoid.

(b) Changing $E$ fields cause $B$ fields. Find the $B$ field (at radius $r$ inside the solenoid) caused by the changing $E$ field that you just found. More precisely, find the difference between the $B$ at radius $r$ and the $B$ on the axis. Label this difference as $\Delta B(r, t)$.

(c) The total $B$ field does not equal $\mu_0 n I_0 \cos \omega t$ throughout the solenoid, due to the $\Delta B(r, t)$ difference you just found. What is the ratio $\Delta B(r, t)/B_0(t)$? Explain why we are justified in making the statement, “The magnetic field inside the solenoid is essentially equal to the naive $\mu_0 n I_0 \cos \omega t$ value, provided that the changes in the current occur on a time scale that is long compared with the time it takes light to travel across the width of the solenoid.” (This time is very short, so for an “everyday” value of $\omega$, the field is essentially equal to $\mu_0 n I_0 \cos \omega t$.)

9.7 Traveling and standing waves **

Consider the two oppositely traveling electric-field waves,

$$E_1 = \hat{x} E_0 \cos(kz - \omega t) \quad \text{and} \quad E_2 = \hat{x} E_0 \cos(kz + \omega t). \quad (9.60)$$

The sum of these two waves is the standing wave, $2\hat{x} E_0 \cos kz \cos \omega t$.

(a) Find the magnetic field associated with this standing electric wave by finding the $B$ fields associated with each of the above traveling $E$ fields, and then adding them.

(b) Find the magnetic field by instead using Maxwell’s equations to find the $B$ field associated with the standing electric wave, $2\hat{x} E_0 \cos kz \cos \omega t$.

9.8 Sunlight *

The power density in sunlight, at the earth, is roughly 1 kilowatt/m$^2$. How large is the rms magnetic field strength?

9.9 Energy flow for a standing wave **

(a) Consider the standing wave in Eq. (9.32). Draw plots of the energy density $U(y, t)$ at $\omega t$ values of $0, \pi/4, \pi/2, 3\pi/4, \text{and } \pi$.

(b) Make a plot of the $y$ component of the Poynting vector, $S_y(y, t)$, at $\omega t$ values of $\pi/4, \pi/2, \text{and } 3\pi/4$. Explain why these plots are consistent with how the energy sloshes back and forth between the different energy plots.

9.10 Energy flow from a wire **

A very thin straight wire carries a constant current $I$ from infinity radially inward to a spherical conducting shell with radius $R$. The

3 Of course, this $\Delta B(r, t)$ difference causes another $E$ field, and so on. So we would get an infinite series of corrections if we kept going. But as long as the current doesn’t change too quickly, the higher-order terms are negligible.
increase in the charge on the shell causes the electric field in the surrounding space to increase, which means that the energy density increases. This implies that there must be a flow of energy from somewhere. This “somewhere” is the wire. Verify that the total flux of the Poynting vector away from a thin tube surrounding the wire equals the rate of change of the energy stored in the electric field. (You can assume that the radius of the wire is much smaller than the radius of the tube, which in turn is much smaller than the radius of the shell.)

9.11 *Momentum in an electromagnetic field** **
We know from Section 9.6 that traveling electromagnetic waves carry energy. But the theory of relativity tells us that anything that transports energy must also transport momentum. Since light may be considered to be made of massless particles (photons), the relation \( p = E/c \) must hold; see Eq. (G.19). In terms of \( E \) and \( B \), find the momentum density of a traveling electromagnetic wave. That is, find the quantity that, when integrated over a given volume, yields the momentum contained in the wave in that volume.

Although we won’t prove it here, the result that you just found for traveling waves is a special case of the more general result that the momentum density equals \( 1/c^2 \) times the energy flow per area per time. This holds for any type of energy flow (matter or field). In particular, it holds for any type of electromagnetic field; even a static field with a nonzero \( E \times B/\mu_0 \) Poynting vector carries momentum. For a nice example of this, see Problem 9.12.

9.12 *Angular momentum paradox***
A setup consists of three very long coaxial cylindrical objects: a nonconducting cylindrical shell with radius \( a \) and total (uniform) charge \( Q \), another nonconducting cylindrical shell with radius \( b > a \) and total (uniform) charge \( -Q \), and a solenoid with radius \( R > b \); see Fig. 9.13. (This setup is a variation of the setup in Boos (1984).)

The current in the solenoid produces a uniform magnetic field \( B_0 \) in its interior. The solenoid is fixed, but the two cylinders are free to rotate (independently) around the axis. They are initially at rest. Imagine that the current in the solenoid is then decreased to zero. (If you want to be picky about keeping the system isolated from external torques, you can imagine the current initially flowing in a superconductor which becomes a normal conductor when heated up.) The changing \( B \) field inside the solenoid will induce an \( E \) field at the locations of the two cylinders.

(a) Find the angular momentum gained by each cylinder by the time the magnetic field has decreased to zero.

(b) You should find that the total change in angular momentum of the cylinders is not zero. Does this mean that angular
momentum isn’t conserved? If it \textit{is} conserved, verify this quantitatively. You may assume that the two cylinders are massive enough so that they don’t end up spinning very quickly, which means that we can ignore the \( B \) fields they generate. \textit{Hint:} See Problem 9.11.

\section*{Exercises}

\subsection*{9.13 \textit{Displacement-current flux} \*}

The flux of the real current through the surface \( S \) in Fig. 9.4 is simply \( I \). Verify explicitly that the flux of the displacement current, \( J_d \equiv \epsilon_0 (\partial E / \partial t) \), through the surface \( S' \) also equals \( I \). What about the sign of the flux? As usual, work in the approximation where the spacing between the capacitor plates is small.

\subsection*{9.14 \textit{Sphere with a hole} \**}

A current \( I \) flows along a wire toward a point charge, causing the charge to increase with time. Consider a spherical surface \( S \) centered at the charge, with a tiny hole where the wire is, as shown in Fig. 9.14. The circumference \( C \) of this hole is the boundary of the surface \( S \). Verify that the integral form of Maxwell’s equation,

\[
\int_C \mathbf{B} \cdot d\mathbf{s} = \int_S \left( \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} \right) \cdot d\mathbf{a}, \quad (9.61)
\]

is satisfied.

\subsection*{9.15 \textit{Field inside a discharging capacitor} \**}

The magnetic field inside the discharging capacitor shown in Fig. 9.1 can in principle be calculated by summing the contributions from all elements of conduction current, as indicated in Fig. 9.5. That might be a long job. If we can assume symmetry about this axis, it is very much easier to find the field \( \mathbf{B} \) at a point by using the integral law,

\[
\int_C \mathbf{B} \cdot d\mathbf{s} = \int_S \left( \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} \right) \cdot d\mathbf{a}, \quad (9.62)
\]

applied to a circular path through the point. Use this to show that the field at \( P \), which is midway between the capacitor plates in Fig. 9.15, and a distance \( r \) from the axis of symmetry, equals \( \mathbf{B} = \mu_0 \mathbf{I} r / 2\pi b^2 \). You may assume that the distance \( s \) between the plates is small compared with their radius \( b \). (Compare this with the calculation of the induced electric field \( \mathbf{E} \) in the example of Fig. 7.16.)

\subsection*{9.16 \textit{Changing flux from a moving charge} \**}

In terms of the electric field \( \mathbf{E} \) of a point charge moving with constant velocity \( \mathbf{v} \), the Lorentz transformation gives the magnetic
Maxwell’s equations and EM waves

\[ \mathbf{B} = \left( \frac{\mathbf{v}}{c^2} \right) \times \mathbf{E} \]

Verify that Maxwell’s equation in integral form, \( \int \mathbf{B} \cdot d\mathbf{s} = \left( \frac{1}{c^2} \right) \left( \frac{d\Phi_E}{dt} \right) \), holds for the circle shown in Fig. 9.16. (We can therefore think of the magnetic field as being induced by the changing electric field of the moving charge.) Hint: Indicate geometrically the new electric flux that passes through the circle after the charge has moved a small distance to the right.

9.17 Gaussian conditions
Start with the source-free, or “empty-space,” Maxwell’s equations in Gaussian units in Eq. (9.21). Consider a wave described by Eqs. (9.22) and (9.23), but now with \( E_0 \) in statvolts/cm and \( B_0 \) in gauss. What conditions must \( E_0, B_0, \) and \( v \) meet to satisfy Maxwell’s equations?

9.18 Associated B field
If the electric field in free space is \( \mathbf{E} = E_0(\hat{x} + \hat{y}) \sin[(2\pi/\lambda)(z + ct)] \) with \( E_0 = 20 \) volts/m, then the magnetic field, not including any static magnetic field, must be what?

9.19 Find the wave
Write out formulas for \( \mathbf{E} \) and \( \mathbf{B} \) that specify a plane electromagnetic sinusoidal wave with the following characteristics. The wave is traveling in the direction \(-\hat{x}\); its frequency is 100 megahertz (MHz), or \( 10^8 \) cycles per second; the electric field is perpendicular to the \( \hat{z} \) direction.

9.20 Kicked by a wave
A free proton was at rest at the origin before the wave described by Eq. (9.28) came past. Let the amplitude \( E_0 \) equal 100 kilovolts/m. Where would you expect to find the proton at time \( t = 1 \) microsecond? The proton mass is \( 1.67 \cdot 10^{-27} \) kg. Hint: Since the duration of the pulse is only a few nanoseconds, you can neglect the displacement of the proton during the passage of the pulse. Also, if the velocity of the proton is not too large, you may ignore the effect of the magnetic field on its motion. The first thing to calculate is the momentum acquired by the proton during the pulse.

9.21 Effect of the magnetic field
Suppose that in Exercise 9.20 the effect of the magnetic field was not entirely negligible. How would it change the direction of the proton’s final velocity? (It suffices to give the dependence on the various parameters; you can ignore any numerical factors.)

9.22 Plane-wave pulse
Consider the plane-wave pulse of \( \mathbf{E} \) and \( \mathbf{B} \) fields shown in Fig. 9.17; \( \mathbf{E} \) points out of the page, and \( \mathbf{B} \) points downward. The fields are uniform inside a “slab” region and are zero outside. The slab has length \( d \) in the \( x \) direction and large (essentially infinite) lengths in the \( y \) and \( z \) directions. It moves with speed \( v \) (to be determined)
in the \( x \) direction. This slab can be considered to be a small section of the transition shell in Appendix H. However, you need not worry about how these fields were generated. All that matters is that the electromagnetic field is self-sustaining via the two “induction” Maxwell equations.

(a) With the dashed rectangular loop shown (which is fixed in space, while the slab moves), use the integral form of one of Maxwell’s equations to obtain a relation between \( E \) and \( B \).

(b) Make a similar argument, with a loop perpendicular to the plane of the page, to obtain another relation between \( E \) and \( B \). (Be careful with the signs.) Then solve for \( v \).

9.23 Field in a box ***

Show that the electromagnetic field described by

\[
E = E_0 \hat{z} \cos kx \cos ky \cos \omega t, \\
B = B_0 (\hat{x} \cos kx \sin ky - \hat{y} \sin kx \cos ky) \sin \omega t
\]  

(9.63)

will satisfy the empty-space Maxwell equations in Eq. (9.18) if \( E_0 = \sqrt{2} c B_0 \) and \( \omega = \sqrt{2} c k \). This field can exist inside a square metal box enclosing the region \(-\pi/2k < x < \pi/2k\) and \(-\pi/2k < y < \pi/2k\), with arbitrary height in the \( z \) direction. Roughly what do the electric and magnetic fields look like?

9.24 Satellite signal *

From a satellite in stationary orbit, a signal is beamed earthward with a power of 10 kilowatts and a beam width covering a region roughly circular and 1000 km in diameter. What is the electric field strength at the receivers, in millivolts/meter?

9.25 Microwave background radiation **

Of all the electromagnetic energy in the universe, by far the largest amount is in the form of waves with wavelengths in the millimeter range. This is the cosmic microwave background radiation discovered by Penzias and Wilson in 1965. It apparently fills all space, including the vast space between galaxies, with an energy density of \( 4 \cdot 10^{-14} \) joule/m\(^3\). Calculate the rms electric field strength in this radiation, in volts/m. Roughly how far away from a 1 kilowatt radio transmitter would you find a comparable electromagnetic wave intensity?

9.26 An electromagnetic wave **

Here is a particular electromagnetic field in free space:

\[
E_x = 0, \quad E_y = E_0 \sin(kx + \omega t), \quad E_z = 0; \\
B_x = 0, \quad B_y = 0, \quad B_z = -(E_0/c) \sin(kx + \omega t).
\]

(9.64)
Maxwell’s equations and EM waves

(a) Show that this field can satisfy Maxwell’s equations if $\omega$ and $k$ are related in a certain way.

(b) Suppose $\omega = 1 \times 10^6 \text{ s}^{-1}$ and $E_0 = 1 \text{ kV/m}$. What is the wavelength? What is the energy density in joules per cubic meter, averaged over a large region? From this calculate the power density, the energy flow in joules per square meter per second.

9.27 Reflected wave  **

A sinusoidal wave is reflected at the surface of a medium whose properties are such that half the incident energy is absorbed. Consider the field that results from the superposition of the incident and the reflected wave. An observer stationed somewhere in this field finds the local electric field oscillating with a certain amplitude $E$. What is the ratio of the largest such amplitude noted by any observer to the smallest amplitude noted by any observer? (This is called the voltage standing wave ratio, or, in laboratory jargon, VSWR.)

9.28 Poynting vector and resistance heating  **

A longitudinal $E$ field inside a wire causes a current via $J = \sigma E$. And since the curl of $E$ is zero, this same longitudinal $E$ component must also exist right outside the surface of the wire. Show that the Poynting vector flux through a cylinder right outside the wire accounts for the $IV$ resistance heating.

9.29 Energy flow in a capacitor  **

A capacitor is charged by having current flow in a thin straight wire from the middle of one circular plate to the middle of the other (as opposed to wires coming in from infinity, as in the example in Section 9.6.2). The electric field inside the capacitor increases, so the energy density also increases. This implies that there must be a flow of energy from somewhere. As in Problem 9.10, this “somewhere” is the wire. Verify that the flux of the Poynting vector away from the wire equals the rate of change of the energy stored in the field. (Of course, we would need to place a battery somewhere along the wire to produce the current flow, and this battery is where the energy flow originates. See Galili and Goihbarg (2005).)

9.30 Comparing the energy densities  **

Consider the capacitor example in Section 9.6.2, but now let the current change in a way that makes the electric field inside the capacitor take the form of $E(t) = E_0 \cos \omega t$. The induced magnetic field is given in Eq. (9.46). Show that the energy density of the magnetic field is much smaller than the energy density of the electric field, provided that the time scale of $\omega$ (namely $2\pi/\omega$) is much longer than the time it takes light to travel across the diameter of the capacitor disks. (As in Problem 9.6, we are ignoring higher-order effects.)
9.31 **Field momentum of a moving charge** ***

Consider a charged particle in the shape of a small spherical shell with radius \( a \) and charge \( q \). It moves with a nonrelativistic speed \( v \). The electric field due to the shell is essentially given by the simple Coulomb field, and the magnetic field is then given by Eq. (6.81). Using the result from Problem 9.11, integrate the momentum density over all space. Show that the resulting total momentum of the electromagnetic field can be written as \( mv \), where \( m \equiv (4/3)(q^2/8\pi\varepsilon_0a)/c^2 \).

An interesting aside: for nonrelativistic speeds, the total energy in the electromagnetic field is dominated by the electric energy. So from Problem 1.32 the total energy in the field equals \( U = (q^2/8\pi\varepsilon_0a) \). Using the above value of \( m \), we therefore find that \( U = (3/4)mc^2 \). This doesn’t agree with Einstein’s \( U = \gamma mc^2 \) result, with \( \gamma \approx 1 \) for a nonrelativistic particle. The qualitative resolution to this puzzle is that, although we correctly calculated the electromagnetic energy, this isn’t the total energy. There must be other forces at play, of course, because otherwise the Coulomb repulsion would cause the particle to fly apart.

9.32 **A Lorentz invariant** ***

Starting from the field transformation given by Eq. (6.76), show that the scalar quantity \( E^2 - c^2B^2 \) is invariant under the transformation. In other words, show that \( E'^2 - c^2B'^2 = E^2 - c^2B^2 \). You can do this using only vector algebra, without writing out \( x, y, z \) components of anything. (The resolution into parallel and perpendicular vectors is convenient for this, since \( E_\perp \cdot E_\parallel = 0 \), \( B_\parallel \times E_\parallel = 0 \), etc.)
Electric fields in matter

Overview In this chapter we study how electric fields affect, and are affected by, matter. We concern ourselves with insulators, or dielectrics, characterized by a dielectric constant. The study of electric fields in matter is largely the study of dipoles. We discussed these earlier in Chapter 2, but we will derive their properties in more generality here, showing in detail how the multipole expansion comes about. The net dipole moment induced in matter by an electric field can come about in two ways. In some cases the electric field polarizes the molecules; the atomic polarizability quantifies this effect. In other cases a molecule has an inherent dipole moment, and the external field serves to align these moments. In any case, a material can be described by a polarization density $P$. The electric susceptibility gives (up to a factor of $\epsilon_0$) the ratio of $P$ to the electric field. The effect of the polarization density is to create a surface charge density on a dielectric material. This explains why the capacitance of a capacitor is increased when it is filled with a dielectric; the surface charge on the dielectric partially cancels the free charge on the capacitor plates.

We study the special case of a uniformly polarized sphere, which interestingly has a uniform electric field in its interior. We then extend this result to the case of a dielectric sphere placed in a uniform electric field. By considering separately the free charge and bound charge, we are led to the electric displacement vector $D$, whose divergence involves only the free charge (unlike the electric field, whose divergence involves all the charge, by Gauss's law). We look at the effects of temperature on the polarization density, how the polarization responds to rapidly changing
fields, and how the bound-charge current affects the “curl $B$” Maxwell equation. Finally, we consider an electromagnetic wave in a dielectric. We find that only a slight modification to the vacuum case is needed.

### 10.1 Dielectrics

The capacitor we studied in Chapter 3 consisted of two conductors, insulated from one another, with nothing in between. The system of two conductors was characterized by a certain capacitance $C$, a constant relating the magnitude of the charge $Q$ on the capacitor (positive charge $Q$ on one plate, equal negative charge on the other) to the difference in electric potential between the two conductors, $\phi_1 - \phi_2$. Let’s denote the potential difference by $\phi_{12}$:

$$C = \frac{Q}{\phi_{12}}. \quad (10.1)$$

For the parallel-plate capacitor, two flat plates each of area $A$ and separated by a distance $s$, we found that the capacitance is given by

$$C = \frac{\epsilon_0 A}{s}. \quad (10.2)$$

Capacitors like this can be found in some electrical apparatus. They are called *vacuum capacitors* and consist of plates enclosed in a highly evacuated bottle. They are used chiefly where extremely high and rapidly varying potentials are involved. Far more common, however, are capacitors in which the space between the plates is filled with some non-conducting solid or liquid substance. Most of the capacitors you have worked with in the laboratory are of that sort; there are dozens of them in any television screen. For conductors embedded in a material medium, Eq. (10.2) does not agree with experiment. Suppose we fill the space between the two plates shown in Fig. 10.1(a) with a slab of plastic, as in Fig. 10.1(b). Experimenting with this new capacitor, we still find a simple proportionality between charge and potential difference, so that we can still define a capacitance by Eq. (10.1). But we find $C$ to be substantially larger than Eq. (10.2) would have predicted. That is, we find more charge on each of the plates, for the same potential difference, plate area, and distance of separation. The plastic slab must be the cause of this.

It is not hard to understand in a general way how this comes about. The plastic slab consists of molecules, the molecules are composed of atoms, which in turn are made of electrically charged particles – electrons and atomic nuclei. The electric field between the capacitor plates acts on those charges, pulling the negative charges up, if the upper plate is positive as in Fig. 10.2, and pushing the positive charges down. Nothing moves very far. (There are no free electrons around, already detached

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure10.1.png}
\caption{(a) A capacitor formed by parallel conducting plates. (b) The same plates with a slab of insulator in between.}
\end{figure}
How a dielectric increases the charge on the plates of a capacitor. (a) Space between the plates empty; \( Q_0 = C_0 \phi_{12} \). (b) Space between the plates filled with a nonconducting material, that is, a dielectric. Electric field pulls negative charges up and pushes positive charges down, exposing a layer of uncompensated negative charge on the upper surface of the dielectric and a layer of uncompensated positive charge on the lower surface. The total charge at the top, including charge \( Q \) on the upper plate, is the same as in (a). \( Q \) itself is now greater than \( Q_0 \); \( Q = \kappa Q_0 \). This \( Q \) is the amount of charge that will flow through the resistor \( R \) if the capacitor is discharged by throwing the switch.

In the presence of the induced layer of negative charge below the upper plate, the charge \( Q \) on the plate itself will increase. In fact, \( Q \) must increase until the total charge at the top, the algebraic sum of \( Q \) and the induced charge layer, equals \( Q_0 \) (the charge on the upper plate before the plastic was inserted). We shall be able to prove this when we return to this problem in Section 10.8 after settling some questions about the electric field inside matter. The important point now is that the charge \( Q \) in Fig. 10.2(b) is larger than \( Q_0 \) and that this \( Q \) is the charge of the capacitor in the relation \( Q = C \phi_{12} \). It is the charge that came out of the battery, and it is the amount of charge that would flow through the resistor \( R \) were we to discharge the capacitor by throwing the switch in the diagram. If we did that, the induced charge layer, which is not part of \( Q \), would simply disappear into the slab.

According to this explanation, the ability of a particular material to increase the capacitance ought to depend on the amount of electric charge in its structure and the ease with which the electrons can be displaced with respect to the atomic nuclei. The factor by which the capacitance is increased when an empty capacitor is filled with a particular material, \( Q/Q_0 \) in our example, is called the dielectric constant of that material. The symbol \( \kappa \) is usually used for it:

\[
Q = \kappa Q_0 \quad \Leftrightarrow \quad C = \kappa C_0
\]

The material itself is often called a dielectric when we are talking about its behavior in an electric field. But any homogeneous nonconducting
Table 10.1.
Dielectric constants of various substances

<table>
<thead>
<tr>
<th>Substance</th>
<th>Conditions</th>
<th>Dielectric constant (( \kappa ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>gas, 0°C, 1 atm</td>
<td>1.00059</td>
</tr>
<tr>
<td>Methane, CH(_4)</td>
<td>gas, 0°C, 1 atm</td>
<td>1.00088</td>
</tr>
<tr>
<td>Hydrogen chloride, HCl</td>
<td>gas, 0°C, 1 atm</td>
<td>1.0046</td>
</tr>
<tr>
<td>Water, H(_2)O</td>
<td>gas, 110°C, 1 atm</td>
<td>1.0126</td>
</tr>
<tr>
<td></td>
<td>liquid, 20°C</td>
<td>80.4</td>
</tr>
<tr>
<td>Benzene, C(_6)H(_6)</td>
<td>liquid, 20°C</td>
<td>2.28</td>
</tr>
<tr>
<td>Methanol, CH(_3)OH</td>
<td>liquid, 20°C</td>
<td>33.6</td>
</tr>
<tr>
<td>Ammonia, NH(_3)</td>
<td>liquid, −34°C</td>
<td>22.6</td>
</tr>
<tr>
<td>Mineral oil</td>
<td>liquid, 20°C</td>
<td>2.24</td>
</tr>
<tr>
<td>Sodium chloride, NaCl</td>
<td>solid, 20°C</td>
<td>6.12</td>
</tr>
<tr>
<td>Sulfur, S</td>
<td>solid, 20°C</td>
<td>4.0</td>
</tr>
<tr>
<td>Silicon, Si</td>
<td>solid, 20°C</td>
<td>11.7</td>
</tr>
<tr>
<td>Polyethylene</td>
<td>solid, 20°C</td>
<td>2.25–2.3</td>
</tr>
<tr>
<td>Porcelain</td>
<td>solid, 20°C</td>
<td>6.0–8.0</td>
</tr>
<tr>
<td>Paraffin wax</td>
<td>solid, 20°C</td>
<td>2.1–2.5</td>
</tr>
<tr>
<td>Pyrex glass 7070</td>
<td>solid, 20°C</td>
<td>4.0</td>
</tr>
</tbody>
</table>

The dielectric constant of a particular material has been determined, perhaps by measuring the capacitance of one capacitor filled with it, we are able to predict the behavior, not merely of two-plate capacitors, but of any electrostatic system made up of conductors and pieces of that substance can be so characterized. Table 10.1 lists the measured values of the dielectric constants for a miscellaneous assortment of substances.

Every dielectric constant in the table is larger than 1. We should expect that if our explanation is correct. The presence of a dielectric could reduce the capacitance below that of the empty capacitor only if its electrons moved, when the electric field was applied, in a direction opposite to the resulting force. For oscillating electric fields, by the way, some such behavior would not be absurd. But for the steady fields we are considering here it can’t work that way.

The dielectric constant of a perfect vacuum is, of course, exactly 1.0 by our definition. For gases under ordinary conditions, \( \kappa \) is only a little larger than 1.0, simply because a gas is mostly empty space. Ordinary solids and liquids usually have dielectric constants ranging from 2 to 6 or so. Note, however, that liquid ammonia is an exception to this rule, and water is a spectacular exception. Actually liquid water is slightly conductive, but that, as we shall have to explain later, does not prevent our defining and measuring its dielectric constant. The ionic conductivity of the liquid is not the reason for the gigantic dielectric constant of water. You can discern this extraordinary property of water in the dielectric constant of the vapor if you remember that it is really the difference between \( \kappa \) and 1 that reveals the electrical influence of the material. Compare the values of \( \kappa \) given in the table for water vapor and for air.

Once the dielectric constant of a particular material has been determined, perhaps by measuring the capacitance of one capacitor filled with it, we are able to predict the behavior, not merely of two-plate capacitors, but of any electrostatic system made up of conductors and pieces of that...
Electric fields in matter

dielectric of any shape. That is, we can predict all electric fields that will exist in the vacuum outside the dielectrics for given charges or potentials on the conductors in the system.

The theory that enables us to do this was fully worked out by the physicists of the nineteenth century. Lacking a complete picture of the atomic structure of matter, they were more or less obliged to adopt a macroscopic description. From that point of view, the interior of a dielectric is a featureless expanse of perfectly smooth “mathematical jelly” whose single electrical property distinguishing it from a vacuum is a dielectric constant different from unity.

If we develop only a macroscopic description of matter in an electric field, we shall find it hard to answer some rather obvious-sounding questions – or, rather, hard to ask these questions in such a way that they can be meaningfully answered. For instance, what is the strength of the electric field inside the plastic slab of Fig. 10.1(b) when there are certain charges on the plates? Electric field strength is defined by the force on a test charge. How can we put a test charge inside a perfectly dense solid, without disturbing anything, and measure the force on it? What would that force mean if we did measure it? You might think of boring a hole and putting the test charge in the hole with some room to move around, so that you can measure the force on it as on a free particle. But then you will be measuring not the electric field in the dielectric, but the electric field in a cavity in the dielectric, which is quite a different thing.

Fortunately another line of attack is available to us, one that leads up from the microscopic or atomic level. We know that matter is made of atoms and molecules; these in turn are composed of elementary charged particles. We know something about the size and structure of these atoms, and we know something about their arrangement in crystals and fluids and gases. Instead of describing our dielectric slab as a volume of structureless but nonvacuous jelly, we shall describe it as a collection of molecules inhabiting a vacuum. If we can find out what the electric charges in one molecule do when that molecule is all by itself in an electric field, we should be able to understand the behavior of two such molecules a certain distance apart in a vacuum. It will only be necessary to include the influence, on each molecule, of any electric field arising from the other. This is a vacuum problem. Now all we have to do is extend this to a population of, say, $10^{20}$ molecules occupying a cubic centimeter or so of vacuum, and we have our real dielectric. We hope to do this without generating $10^{20}$ separate problems.

This program, if carried through, will reward us in two ways. We shall be able at last to say something meaningful about the electric and magnetic fields inside matter, answering questions such as the one raised above. What is more valuable, we shall understand how the macroscopic electric and magnetic phenomena in matter arise from, and therefore reveal, the nature of the underlying atomic structure. We are going to study electric and magnetic effects separately. We begin with dielectrics.
Since our first goal is to describe the electric field produced by an atom or molecule, it will help to make some general observations about the electrostatic field external to any small system of charges.

10.2 The moments of a charge distribution

An atom or molecule consists of some electric charges occupying a small volume, perhaps a few cubic angstroms \((10^{-30} \text{ m}^3)\) of space. We are interested in the electric field outside that volume, which arises from this rather complicated charge distribution. We shall be particularly concerned with the field far away from the source, by which we mean far away compared with the size of the source itself. What features of the charge structure mainly determine the field at remote points? To answer this, let’s look at some arbitrary distribution of charges and see how we might go about computing the field at a point outside it. The discussion in this and the following section generalizes our earlier discussion of dipoles in Section 2.7.

Figure 10.3 shows a charge distribution of some sort located in the neighborhood of the origin of coordinates. It might be a molecule consisting of several positive nuclei and quite a large number of electrons. In any case we shall suppose it is described by a given charge density function \(\rho(x,y,z)\); \(\rho\) is negative where the electrons are and positive where the nuclei are. To find the electric field at distant points we can begin by computing the potential of the charge distribution. To illustrate, let’s take some point \(A\) out on the \(z\) axis. (Since we are not assuming any special symmetry in the charge distribution, there is nothing special about the \(z\) axis.) Let \(r\) be the distance of \(A\) from the origin. The electric potential at \(A\), denoted by \(\phi_A\), is obtained as usual by adding the contributions from all elements of the charge distribution:

\[
\phi_A = \frac{1}{4\pi \epsilon_0} \int \frac{\rho(x',y',z') \, dv'}{R}. \tag{10.4}
\]

In the integrand, \(dv'\) is an element of volume within the charge distribution, \(\rho(x',y',z')\) is the charge density there, and \(R\) in the denominator is the distance from \(A\) to this particular charge element. The integration is carried out in the coordinates \(x',y',z'\), of course, and is extended over all the region containing charge. We can express \(R\) in terms of \(r\) and the distance \(r'\) from the origin to the charge element. Using the law of cosines with \(\theta\) the angle between \(r'\) and the axis on which \(A\) lies, we have

\[
R = \left(r^2 + r'^2 - 2rr' \cos \theta \right)^{1/2}. \tag{10.5}
\]

With this substitution for \(R\), the integral becomes

\[
\phi_A = \frac{1}{4\pi \epsilon_0} \int \rho \, dv' \left(r^2 + r'^2 - 2rr' \cos \theta \right)^{-1/2}. \tag{10.6}
\]
Now we want to take advantage of the fact that, for a distant point like $A$, $r'$ is much smaller than $r$ for all parts of the charge distribution. This suggests that we should expand the square root in Eq. (10.5) in powers of $r'/r$. Writing
\[
(r^2 + r'^2 - 2rr' \cos \theta)^{-1/2} = \frac{1}{r} \left[ 1 + \left( \frac{r'^2}{r^2} - \frac{2r'}{r} \cos \theta \right) \right]^{-1/2}
\] (10.7)
and using the expansion $(1 + \delta)^{-1/2} = 1 - \delta/2 + 3\delta^2/8 - \cdots$, we get, after collecting together terms of the same power in $r'/r$, the following:
\[
(r^2 + r'^2 - 2rr' \cos \theta)^{-1/2} = \frac{1}{r} \left[ 1 + \frac{r'}{r} \cos \theta + \left( \frac{r'}{r} \right)^2 \frac{(3\cos^2 \theta - 1)}{2} + O \left( \left( \frac{r'}{r} \right)^3 \right) \right],
\] (10.8)
where the last term here indicates terms of order at least $(r'/r)^3$. These are very small if $r' \ll r$. Now, $r$ is a constant in the integration, so we can take it outside and write the prescription for the potential at $A$ as follows:
\[
\phi_A = \frac{1}{4\pi \varepsilon_0} \left[ \frac{1}{r} \int_{K_0} \rho \, dv' + \frac{1}{r^2} \int_{K_1} r' \cos \theta \, \rho \, dv' \right] + \frac{1}{r^3} \int_{K_2} r'^2 \left( \frac{3\cos^2 \theta - 1}{2} \right) \rho \, dv' + \cdots.
\] (10.9)
Each of the integrals above, $K_0$, $K_1$, $K_2$, and so on, has a value that depends only on the structure of the charge distribution, not on the distance to point $A$. Hence the potential for all points along the $z$ axis can be written as a power series in $1/r$ with constant coefficients:
\[
\phi_A = \frac{1}{4\pi \varepsilon_0} \left[ \frac{K_0}{r} + \frac{K_1}{r^2} + \frac{K_2}{r^3} + \cdots \right].
\] (10.10)
This power series is called the multipole expansion of the potential, although we have calculated it only for a point on the $z$ axis here. To finish the problem we would have to get the potential $\phi$ at all other points, in order to calculate the electric field as $-\text{grad} \, \phi$. We have gone far enough, though, to bring out the essential point: The behavior of the potential at large distances from the source will be dominated by the first term in the above series whose coefficient is not zero.

Let us look at these coefficients more closely. The coefficient $K_0$ is $\int \rho \, dv'$, which is simply the total charge in the distribution. If we have equal amounts of positive and negative charge, as in a neutral molecule,
$K_0$ will be zero. For a singly ionized molecule, $K_0$ will have the value $e$. If $K_0$ is not zero, then no matter how large $K_1$, $K_2$, etc., may be, if we go out to a sufficiently large distance, the term $K_0/r$ will win out. Beyond that, the potential will approach that of a point charge at the origin and so will the field. This is hardly surprising.

Suppose we have a neutral molecule, so that $K_0$ is equal to zero. Our interest now shifts to the second term, with coefficient $K_1 = \int r' \cos \theta \rho \, dv'$. Since $r' \cos \theta$ is simply $z'$, this term measures the relative displacement, in the direction toward $A$, of the positive and negative charge. It has a nonzero value for the distributions sketched in Fig. 10.4, where the densities of positive and negative charge have been indicated separately. In fact, all the distributions shown have approximately the same value of $K_1$. Furthermore – and this is a crucial point – if any charge distribution is neutral, the value of $K_1$ is independent of the position chosen as origin. That is, if we replace $z'$ by $z' + z_0'$, in effect shifting the origin, the value of the integral is not changed: $\int (z' + z_0') \rho \, dv' = \int z' \rho \, dv' + z_0' \int \rho \, dv'$, and the latter integral is always zero for a neutral distribution.

Evidently, if $K_0 = 0$ and $K_1 \neq 0$, the potential along the $z$ axis will vary asymptotically (that is, with ever-closer approximation as we go out to larger distances) as $1/r^2$. We recognize this dependence on $r$ from the dipole discussion in Section 2.7. We expect the electric field strength to behave asymptotically like $1/r^3$, in contrast with the $1/r^2$ dependence of the field from a point charge. Of course, we have discussed only the potential on the $z$ axis. We will return to the question of the exact form of the field after getting a general view of the situation.

If $K_0$ and $K_1$ are both zero, and $K_2$ is not, the potential will behave like $1/r^3$ at large distances, and the field strength will fall off with the inverse fourth power of the distance. Figure 10.5 shows a charge distribution for which $K_0$ and $K_1$ are both zero (and would be zero no matter what direction we had chosen for the $z$ axis), while $K_2$ is not zero.

The quantities $K_0$, $K_1$, $K_2$, ... are related to what are called the moments of the charge distribution. Using this language, we call $K_0$, which is simply the net charge, the monopole moment, or monopole strength. $K_1$ is one component of the dipole moment of the distribution. The dipole moment has the dimensions (charge) $\times$ (displacement); it is a vector, and our $K_1$ is its $z$ component. The third constant $K_2$ is related to the quadrupole moment of the distribution, the next to the octupole moment, and so on. The quadrupole moment is not a vector, but a tensor. The charge distribution shown in Fig. 10.5 has a nonzero quadrupole moment. You can quickly show that $K_2 = 3e a^2$, where $a$ is the distance from each charge to the origin.

**Figure 10.4.**
Some charge distributions with $K_0 = 0$, $K_1 \neq 0$. That is, each has net charge zero, but nonzero dipole moment.
Electric fields in matter

**Example (Sphere monopole)** The external potential due to a spherical shell with uniform surface charge density is \( Q/4\pi\varepsilon_0 r \). Therefore, its only nonzero moment is the monopole moment. That is, all of the \( K_i \) terms except \( K_0 \) in Eq. (10.10) are zero. Using the integral forms given in Eq. (10.9), verify that \( K_1 \) and \( K_2 \) are zero.

**Solution** For a surface charge density, the \( \rho dv' \) in the \( K_i \) integrals turns into \( \sigma da' = \sigma(2\pi R \sin \theta)(R d\theta) \). Since we’re trying to show that the integrals are zero, the various constants in \( \sigma da' \) don’t matter. Only the angular dependence, \( \sin \theta d\theta \), is relevant. So we have

\[
K_1 \propto \int_0^\pi \cos \theta \sin \theta d\theta = -\frac{1}{2} \cos^2 \theta \bigg|_0^\pi = 0,
\]

\[
K_2 \propto \int_0^\pi (3\cos^2 \theta - 1) \sin \theta d\theta = (-\cos^3 \theta + \cos \theta) \bigg|_0^\pi = 0, \quad (10.11)
\]

as desired. Intuitively, it is clear from symmetry that \( K_1 \) is zero; for every bit of charge with height \( z' \), there is a corresponding bit of charge with height \(-z'\). But it isn’t as intuitively obvious that \( K_2 \) vanishes.

As mentioned above, \( K_1 \) and \( K_2 \) are only components of the complete dipole vector and quadrupole tensor. But the other components can likewise be shown to equal zero, as we know they must. If you want to calculate the general form of the complete quadrupole tensor, one way is to write the \( R \) in Eq. (10.5) as \( R = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2} \), and then perform a Taylor expansion as we did above. See Problem 10.6.

The advantage to us of describing a charge distribution by this hierarchy of moments is that it singles out just those features of the charge distribution that determine the field at a great distance. If we were concerned only with the field in the immediate neighborhood of the distribution, it would be a fruitless exercise. For our main task, understanding what goes on in a dielectric, it turns out that only the monopole strength (the net charge) and the dipole strength of the molecular building blocks are important. We can ignore all other moments. And if the building blocks are neutral, we have only their dipole moments to consider.

### 10.3 The potential and field of a dipole

The dipole contribution to the potential at the point \( A \), at distance \( r \) from the origin, is given by \( (1/4\pi\varepsilon_0 r^2) \int r' \cos \theta \rho dv' \). We can write \( r' \cos \theta \), which is just the projection of \( \mathbf{r}' \) on the direction toward \( A \), as \( \hat{\mathbf{r}} \cdot \mathbf{r}' \). Thus we can write the potential without reference to any arbitrary axis as

\[
\phi_A = \frac{1}{4\pi\varepsilon_0 r^2} \int \hat{\mathbf{r}} \cdot \mathbf{r}' \rho dv' = \frac{\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2} \cdot \int \mathbf{r}' \rho dv', \quad (10.12)
\]

which will serve to give the potential at any point with location \( r\hat{\mathbf{r}} \). The integral on the right in Eq. (10.12) is the **dipole moment** of the charge...
distribution. It is a vector, obviously, with the dimensions (charge) × (distance). We shall denote the dipole moment vector by \( \mathbf{p} \):

\[
\mathbf{p} = \int r' \rho \, dv'
\]

(10.13)

The dipole moment \( p = q \ell \) in Section 2.7 is a special case of this result. If we have two point charges \( \pm q \) located at positions \( z = \pm \ell/2 \), then \( \rho \) is nonzero only at these two points. So the integral in Eq. (10.13) becomes a discrete sum: \( \mathbf{p} = q(\hat{\mathbf{z}} \ell/2) + (-q)(-\hat{\mathbf{z}} \ell/2) = (q \ell) \hat{\mathbf{z}} \), which agrees with the \( p = q \ell \) result in Eq. (2.35). The dipole vector points in the direction from the negative charge to the positive charge.

Using the dipole moment \( \mathbf{p} \), we can rewrite Eq. (10.12) as

\[
\phi(\mathbf{r}) = \frac{\hat{\mathbf{r}} \cdot \mathbf{p}}{4\pi \epsilon_0 r^2}.
\]

(10.14)

The electric field is the negative gradient of this potential. To see what the dipole field is like, locate a dipole \( \mathbf{p} \) at the origin, pointing in the \( z \) direction (Fig. 10.6). With this arrangement,

\[
\phi = \frac{p \cos \theta}{4\pi \epsilon_0 r^2}
\]

(10.15)
Electric fields in matter

in agreement with the result in Eq. (2.35).\(^1\) The potential and the field are, of course, symmetrical around the \(z\) axis. Let’s work with Cartesian coordinates in the \(xz\) plane, where \(\cos \theta = z/(x^2 + z^2)^{1/2}\). In that plane,

\[
\phi = \frac{pz}{4\pi \epsilon_0 (x^2 + z^2)^{3/2}}.
\]  

(10.16)

The components of the electric field are readily derived:

\[
E_x = -\frac{\partial \phi}{\partial x} = \frac{3pxz}{4\pi \epsilon_0 (x^2 + z^2)^{5/2}} = \frac{3p \sin \theta \cos \theta}{4\pi \epsilon_0 r^3},
\]  

(10.17)

\[
E_z = -\frac{\partial \phi}{\partial z} = \frac{p}{4\pi \epsilon_0} \left[ \frac{3z^2}{(x^2 + z^2)^{5/2}} - \frac{1}{(x^2 + z^2)^{3/2}} \right] = \frac{p(3 \cos^2 \theta - 1)}{4\pi \epsilon_0 r^3}.
\]

The dipole field can be described more simply in the polar coordinates \(r\) and \(\theta\). Let \(E_r\) be the component of \(E\) in the direction of \(\hat{r}\), and let \(E_\theta\) be the component perpendicular to \(\hat{r}\) in the direction of increasing \(\theta\). You can show in Problem 10.4 that Eq. (10.17) implies

\[
E_r = \frac{p}{2\pi \epsilon_0 r^3} \cos \theta, \quad E_\theta = \frac{p}{4\pi \epsilon_0 r^3} \sin \theta,
\]  

(10.18)

in agreement with the result in Eq. (2.36). Alternatively, you can quickly derive Eq. (10.18) directly by working in polar coordinates and taking the negative gradient of the potential given by Eq. (10.15). This is the route we took in Section 2.7.

Proceeding out in any direction from the dipole, we find the electric field strength falling off as \(1/r^3\), as we had anticipated. Along the \(z\) axis the field is parallel to the dipole moment \(p\), with magnitude \(p/2\pi \epsilon_0 r^3\); that is, it has the value \(p/(2\pi \epsilon_0 r^3)\). In the equatorial plane the field points antiparallel to \(p\) and has the value \(-p/(4\pi \epsilon_0 r^3)\). This field may remind you of the field in the setup with a point charge over a conducting plane, with its image charge, from Section 3.4. That of course is just the two-charge dipole we discussed in Section 2.7. In Fig. 10.7 we show the field of this pair of charges, mainly to emphasize that the field near the charges is not a dipole field. This charge distribution has many multipole moments, indeed infinitely many, so it is only the far field at distances \(r \gg s\) that can be represented as a dipole field.

To generate a complete dipole field right into the origin we would have to let \(s\) shrink to zero while increasing \(q\) without limit so as to keep \(p = qs\) finite. This highly singular abstraction is not very interesting. We know that our molecular charge distribution will have complicated near fields, so we could not easily represent the near region in any case. Fortunately we shall not need to.

\(^1\) Note that the angle \(\theta\) here has a different meaning from the angle \(\theta\) in Fig. 10.3 and Eqs. (10.5)–(10.9), where it indicated the position of a point in the charge distribution. The present \(\theta\) indicates the position of a given point (at which we want to calculate \(\phi\) and \(E\)) with respect to the dipole direction.
10.4 The torque and the force on a dipole in an external field

Suppose two charges, $q$ and $-q$, are mechanically connected so that $s$, the distance between them, is fixed. You may think of the charges as stuck on the end of a short nonconducting rod of length $s$. We shall call this object a dipole. Its dipole moment $p$ is simply $qs$. Let us put the dipole in an external electric field, that is, the field from some other source. The field of the dipole itself does not concern us now. Consider first a uniform electric field, as in Fig. 10.8(a). The positive end of the dipole is pulled toward the right, the negative end toward the left, by a force of strength $qE$. The net force on the object is zero, and so is the torque, in this position.

A dipole that makes some angle $\theta$ with the field direction, as in Fig. 10.8(b), obviously experiences a torque. In general, the torque $\mathbf{N}$ around an axis through some chosen origin is $\mathbf{r} \times \mathbf{F}$, where $\mathbf{F}$ is the force applied at a position $\mathbf{r}$ relative to the origin. Taking the origin in the center of the dipole, so that $r = s/2$, we have

$$\mathbf{N} = \mathbf{r} \times \mathbf{F}_+ + (-\mathbf{r}) \times \mathbf{F}_-. \quad (10.19)$$

$\mathbf{N}$ is a vector perpendicular to the figure, and its magnitude is given by

$$N = \frac{s}{2} qE \sin \theta + \frac{s}{2} qE \sin \theta = sqE \sin \theta = pE \sin \theta. \quad (10.20)$$

This can be written simply as

$$\mathbf{N} = \mathbf{p} \times \mathbf{E} \quad (10.21)$$

When the total force on the dipole is zero, as it is in this case, the torque is independent of the choice of origin (as you should verify), which therefore need not be specified.
The orientation of the dipole in Fig. 10.8(a) has the lowest energy. Work has to be done to rotate it into any other position. Let us calculate the work required to rotate the dipole from a position parallel to the field, through some angle $\theta_0$, as shown in Fig. 10.8(c). Rotation through an infinitesimal angle $d\theta$ requires an amount of work $N d\theta$. Thus the total work done is

$$\int_0^{\theta_0} N d\theta = \int_0^{\theta_0} pE \sin \theta d\theta = pE(1 - \cos \theta_0).$$  \hspace{1cm} (10.22)

This makes sense, because each charge moves a distance $(s/2)(1 - \cos \theta_0)$ against the field. The force is $qE$, so the work done on each charge is $(qE)(s/2)(1 - \cos \theta_0)$. Doubling this gives the result in Eq. (10.22). To reverse the dipole, turning it end over end, corresponds to $\theta_0 = \pi$ and requires an amount of work equal to $2pE$.

The net force on the dipole in any uniform field is zero, obviously, regardless of its orientation. In a nonuniform field the forces on the two ends of the dipole will generally not be exactly equal and opposite, and there will be a net force on the object. A simple example is a dipole in the field of a point charge $Q$. If the dipole is oriented radially, as in Fig. 10.9(a), with the positive end nearer the positive charge $Q$, the net force will be outward, and its magnitude will be

$$F = (q) \frac{Q}{4\pi \varepsilon_0 r^2} + (-q) \frac{Q}{4\pi \varepsilon_0 (r + s)^2}. \hspace{1cm} (10.23)$$

For $s \ll r$, we need only evaluate this to first order in $s/r$:

$$F \approx \frac{qQ}{4\pi \varepsilon_0 r^2} \left[ 1 - \frac{1}{1 + \frac{s}{r}} \right] \approx \frac{qQ}{4\pi \varepsilon_0 r^2} \left[ 1 - \frac{1 - \frac{2s}{r}}{1 + \frac{2s}{r}} \right] = \frac{sqQ}{2\pi \varepsilon_0 r^3}. \hspace{1cm} (10.24)$$

In terms of the dipole moment $p$, this is simply

$$F = \frac{pQ}{2\pi \varepsilon_0 r^3}. \hspace{1cm} (10.25)$$

With the dipole at right angles to the field, as in Fig. 10.9(b), there is also a force. Now the forces on the two ends, though equal in magnitude, are not exactly opposite in direction. In this case there is a net upward force.

It is not hard to work out a general formula for the force on a dipole in a nonuniform electric field. The force depends essentially on the gradients of the various components of the field. In general, the $x$ component of the force on a dipole of moment $p$ is

$$F_x = p \cdot \text{grad} E_x$$  \hspace{1cm} (10.26)

with corresponding formulas for $F_y$ and $F_z$; see Problem 10.7. All three components can be collected into the concise statement, $\mathbf{F} = (\mathbf{p} \cdot \nabla)\mathbf{E}$. 

---

**Figure 10.8.**
(a) A dipole in a uniform field. (b) The torque on the dipole is $\mathbf{N} = \mathbf{p} \times \mathbf{E}$; the vector $\mathbf{N}$ points into the page. (c) The work done in turning the dipole from an orientation parallel to the field to the orientation shown is $pE(1 - \cos \theta_0)$.
10.5 Atomic and molecular dipoles; induced dipole moments

Consider the simplest atom, the hydrogen atom, which consists of a nucleus and one electron. If you imagine the negatively charged electron revolving around the positive nucleus like a planet around the sun—as in the original atomic model of Niels Bohr—you will conclude that the atom has, at any one instant of time, an electric dipole moment. The dipole moment vector $\mathbf{p}$ points parallel to the electron–proton radius vector, and its magnitude is $e$ times the electron–proton distance. The direction of this vector will be continually changing as the electron, in this picture of the atom, circles around its orbit. To be sure, the time average of $\mathbf{p}$ will be zero for a circular orbit, but we should expect the periodically changing dipole moment components to generate rapidly oscillating electric fields and electromagnetic radiation.

The absence of such radiation in the normal hydrogen atom was one of the baffling paradoxes of early quantum physics. Modern quantum mechanics tells us that it is better to think of the hydrogen atom in its lowest energy state (the usual condition of most of the hydrogen atoms in the universe) as a spherically symmetrical structure with the electronic charge distributed, in the time average, over a cloud surrounding the nucleus. Nothing is revolving in a circle or oscillating. If we could take a snapshot with an exposure time shorter than $10^{-16}$ s, we might discern an electron localized some distance away from the nucleus. But for processes involving times much longer than that, we have, in effect, a smooth distribution of negative charge surrounding the nucleus and extending out in all directions with steadily decreasing density. The total charge in this distribution is just $-e$, the charge of one electron. Roughly half of it lies within a sphere of radius 0.5 angstrom ($0.5 \times 10^{-10}$ m). The density decreases exponentially outward; a sphere only 2.2 angstroms in radius contains 99 percent of the charge. The electric field in the atom is just what a stationary charge distribution of this form, together with the positive nucleus, would produce.

A similar picture is the best one to adopt for other atoms and molecules. We can treat the nuclei in molecules as point charges; for our present purposes their size is too small to matter. The entire electronic structure of the molecule is to be pictured as a single cloud of negative charge of smoothly varying density. The shape of this cloud, and the variation of charge density within it, will of course be different for different molecules. But at the fringes of the cloud the density will always fall off exponentially, so that it makes some sense to talk of the size and shape of the molecular charge distribution.

Quantum mechanics makes a crucial distinction between stationary states and time-dependent states of an atom. The state of lowest energy is a time-independent structure, a stationary state. It has to be, according to the laws of quantum mechanics. It is that state of the atom or...
molecule that concerns us here. Of course, atoms can radiate electromagnetic energy. That happens with the atom in a nonstationary state in which there is an oscillating electric charge.

**Figure 10.10** represents the charge distribution in the normal hydrogen atom. It is a cross section through the spherically symmetrical cloud, with the density suggested by shading. Obviously the dipole moment of such a distribution is zero. The same is true of any atom in its state of lowest energy, no matter how many electrons it contains, for in all such states the electron distribution has spherical symmetry. It is also true of any ionized atom, though an ion of course has a monopole moment, that is, a net charge.

So far we have found nothing very interesting. But now let us put the hydrogen atom in an electric field supplied by some external source, as in **Fig. 10.11**. The electric field distorts the atom, pulling the negative charge down and pushing the positive nucleus up. The distorted atom will have an electric dipole moment because the “center of gravity” of the negative charge will no longer coincide with the positive nucleus, but will be displaced from the nucleus by some small distance \( \Delta z \). The electric dipole moment of the atom is now \( e \Delta z \).

How much distortion will be caused by a field of given strength \( E \)? Remember that electric fields already exist in the unperturbed atom, of strength \( e/4\pi\epsilon_0a^2 \) in order of magnitude, where \( a \) is a typical atomic dimension. We should expect the relative distortion of the atom’s structure, measured by the ratio \( \Delta z/a \), to have the same order of magnitude as the ratio of the perturbing field \( E \) to the internal fields that hold the atom together. We predict, in other words, that

\[
\frac{\Delta z}{a} \approx \frac{E}{e/4\pi\epsilon_0a^2}. \quad (10.27)
\]

If you don’t trust this reasoning, Exercise 10.30 gives an alternative method for finding the relation between \( \Delta z \) and \( E \).

Now \( a \) is a length of order \( 10^{-10} \) m, and \( e/4\pi\epsilon_0a^2 \) is approximately \( 10^{11} \) volts/m, a field thousands of times more intense than any large-scale steady field we could make in the laboratory. Evidently the distortion of the atom is going to be very slight indeed, in any practical case. If Eq. (10.27) is correct, it follows that the dipole moment \( p \) of the distorted atom, which is just \( e \Delta z \), will be

\[
p = e \Delta z \approx 4\pi\epsilon_0a^3 E. \quad (10.28)
\]

Since the atom was spherically symmetrical before the field \( E \) was applied, the dipole moment vector \( p \) will be in the direction of \( E \). The factor that relates \( p \) to \( E \) is called the **atomic polarizability**, and is usually denoted by \( \alpha \):

\[
p = \alpha E \quad (10.29)
\]
It is common to work instead with the quantity $\alpha/4\pi\epsilon_0$, which has the dimensions of volume. The reason for this is that a direct comparison between $p$ and $E$ isn’t quite a fair one, because electric fields contain a somewhat arbitrary factor of $1/4\pi\epsilon_0$ multiplying the factors of charge and distance in Coulomb’s law. A more reasonable comparison would therefore involve $p$ and $4\pi\epsilon_0E$. These quantities have dimensions of (charge) × (distance) and (charge)/(distance)$^2$, respectively. Equation (10.29) then yields $p/(4\pi\epsilon_0E) = \alpha/4\pi\epsilon_0$. This quantity is often also called the atomic polarizability, so the term is a little ambiguous. It is best to say explicitly whether you are working with $\alpha$ or $\alpha/4\pi\epsilon_0$.

According to our estimate in Eq. (10.28), we have $\alpha \approx 4\pi\epsilon_0a^3$, so $\alpha/4\pi\epsilon_0$ is in order of magnitude an atomic volume, something like $a^3 \approx 10^{-30}$ m$^3$. Its value for a particular atom will depend on the details of the atom’s electronic structure. An exact quantum-mechanical calculation of the polarizability of the hydrogen atom predicts $\alpha/4\pi\epsilon_0 = (9/2)a_0^3$, where $a_0$ is the Bohr radius, $0.52 \cdot 10^{-10}$ m, the characteristic distance in the H-atom structure in its normal state. The values of $\alpha/4\pi\epsilon_0$ for several species of atoms, experimentally determined, are given in Table 10.2. The examples given are arranged in order of increasing number of electrons. Note the wide variations in $\alpha/4\pi\epsilon_0$. If you are acquainted with the periodic table of the elements, you may discern something systematic here. Hydrogen and the alkali metals lithium, sodium, and potassium, which occupy the first column of the periodic table, have large values of $\alpha/4\pi\epsilon_0$, and these increase steadily with increasing atomic number, from hydrogen to potassium. The noble gases have much smaller atomic polarizabilities, but these also increase as we proceed, within the family, from helium to neon to krypton. Apparently the alkali atoms, as a class, are easily deformed by an electric field, whereas the electronic structure of a noble gas atom is much stiffer. It is the loosely bound outer, or “valence,” electron in the alkali atom structure that is responsible for the easy polarizability.

A molecule, too, develops an induced dipole moment when an electric field is applied to it. The methane molecule depicted in Fig. 10.12 is made from four hydrogen atoms arranged at the corners of a tetrahedron around the central carbon atom. This object has an electrical polarizability, determined experimentally, of

$$\frac{\alpha}{4\pi\epsilon_0} = 2.6 \cdot 10^{-30} \text{ m}^3.$$  

(10.30)
It is interesting to compare this with the sum of the polarizabilities of a carbon atom and four isolated hydrogen atoms. Taking the data from Table 10.2, we find \( \frac{\alpha_C}{4\pi\varepsilon_0} + 4\frac{\alpha_H}{4\pi\varepsilon_0} = 4.1 \times 10^{-30} \text{ m}^3 \). Evidently the binding of the atoms into a molecule has somewhat altered the electronic structure. Measurements of atomic and molecular polarizabilities have long been used by chemists as clues to molecular structure.

### 10.6 Permanent dipole moments

Some molecules are so constructed that they have electric dipole moments even in the absence of an electric field. They are unsymmetrical in their normal state. The molecule shown in Fig. 10.13 is an example. A simpler example is provided by any diatomic molecule made out of dissimilar atoms, such as hydrogen chloride, H\(_2\)Cl. There is no point on the axis of this molecule about which the molecule is symmetrical fore and aft; the two ends of the molecule are physically different. It would be a pure accident if the center of gravity of the positive charge and that of the negative charge happened to fall at the same point along the axis. When the H\(_2\)Cl molecule is formed from the originally spherical H and Cl atoms, the electron of the H atom shifts partially over to the Cl structure, leaving the hydrogen nucleus partially denuded. So there is some excess of positive charge at the hydrogen end of the molecule and a corresponding excess of negative charge at the chlorine end. The magnitude of the resulting electric dipole moment, \( p = 3.4 \times 10^{-30} \text{ coulomb-meter} \), is equivalent to shifting one electron about 0.2 angstrom (using \( s = p/e \)).

By contrast, the hydrogen atom in a field of 1 megavolt per meter, with the polarizability listed in Table 10.2, acquires an induced moment less than \( 10^{-34} \text{ coulomb-meter} \). Permanent dipole moments, when they exist, are as a rule enormously larger than any moment that can be induced by ordinary laboratory electric fields.\(^2\) Because of this, the distinction between polar molecules, as molecules with “built-in” dipole moments are called, and nonpolar molecules is very sharp.

We said at the beginning of Section 10.5 that the hydrogen atom had, at any instant of time, a dipole moment. But then we dismissed it as being zero in the time average, on account of the rapid motion of the electron. Now we seem to be talking about molecular dipole moments as if a molecule were an ordinary stationary object like a baseball bat whose ends could be examined at leisure to see which was larger! Molecules move more slowly than electrons, but their motion is rapid by ordinary standards. Why can we credit them with “permanent” electric dipole moments? If this inconsistency was bothering you, you are to be commended. The full answer can’t be given without some quantum

---

\(^2\) There is a good reason for this. The internal electric fields in atoms and molecules, as we remarked in Section 10.5, are naturally on the order of \( e/4\pi\varepsilon_0(10^{-10} \text{ m})^2 \), which is roughly \( 10^{11} \text{ volts/m} \)! We cannot apply such a field to matter in the laboratory for the closely related reason that it would tear the matter to bits.
mechanics, but the difference essentially involves the time scale of the motion. The time it takes a molecule to interact with its surroundings is generally shorter than the time it takes the intrinsic motion of the molecule to average out the dipole moment smoothly. Hence the molecule really acts as if it had the moment we have been talking about. A very short time qualifies as permanent in the world of one molecule and its neighbors.

Some common polar molecules are shown in Fig. 10.14, with the direction and magnitude of the permanent dipole moment indicated for each. The water molecule has an electric dipole moment because it is bent in the middle, the O–H axes making an angle of about 105° with one another. This is a structural oddity with the most far-reaching consequences. The dipole moment of the molecule is largely responsible for the properties of water as a solvent, and it plays a decisive role in chemistry that goes on in an aqueous environment. It is hard to imagine what the world would be like if the H₂O molecule, like the CO₂ molecule, had its parts arranged in a straight line; probably we wouldn’t be here to observe it. We hasten to add that the shape of the H₂O molecule is not a capricious whim of Nature. Quantum mechanics has revealed clearly why a molecule made of an eight-electron atom joined to two one-electron atoms must prefer to be bent.

The behavior of a polar substance as a dielectric is strikingly different from that of material composed of nonpolar molecules. The dielectric constant of water is about 80, that of methyl alcohol 33, while a typical nonpolar liquid might have a dielectric constant around 2. In a nonpolar substance the application of an electric field induces a slight dipole moment in each molecule. In the polar substance dipoles are already present in great strength but, in the absence of a field, are pointing in random directions so that they have no large-scale effect. An applied electric field merely aligns them to a certain degree. In either process, however, the macroscopic effects will be determined by the net amount of polarization per unit volume.

### 10.7 The electric field caused by polarized matter

#### 10.7.1 The field outside matter

Suppose we build up a block of matter by assembling a very large number of molecules in a previously empty region of space. Suppose too that each of these molecules is polarized in the same direction. For the present we need not concern ourselves with the nature of the molecules or with the means by which their polarization is maintained. We are

---

**Figure 10.14.**
Some well-known polar molecules. The observed value of the permanent dipole moment $p$ is given in units of $10^{-30}$ coulomb-meters.
interested only in the electric field they produce when they are in this condition; later we can introduce any fields from other sources that might be around. If you like, you can imagine that these are molecules with permanent dipole moments that have been lined up neatly, all pointing the same way, and frozen in position. All we need to specify is \( N \), the number of dipoles per cubic meter, and the moment of each dipole \( p \). We shall assume that \( N \) is so large that any macroscopically small volume \( dv \) contains quite a large number of dipoles. The total dipole strength in such a volume is \( pN \ dv \). At any point far away from this volume element compared with its size, the electric field from these particular dipoles would be practically the same if they were replaced by a single dipole moment of strength \( pN \ dv \). We shall call \( pN \) the density of polarization, and denote it by \( P \), a vector quantity with the units C-m/m\(^3\) (or C/m\(^2\)):

\[
P \equiv pN = \text{dipole moment/volume}.
\] (10.31)

\( P \ dv \) is the dipole moment to be associated with any small-volume element \( dv \) for the purpose of computing the electric field at a distance. By the way, our matter has been assembled from neutral molecules only; there is no net charge in the system or on any molecule, so we have only the dipole moments to consider as sources of a distant field.

Figure 10.15 shows a slender column, or cylinder, of this polarized material. Its cross section is \( da \), and it extends vertically from \( z_1 \) to \( z_2 \). The polarization density \( P \) within the column is uniform over the length and points in the positive \( z \) direction. We are about to calculate the electric potential, at some external point, due to this column of polarization. An element of the cylinder, of height \( dz \), has a dipole moment \( P \ dv \ = \ P \ da \ dz \). Its contribution to the potential at the point \( A \) can be written down by referring back to our formula Eq. (10.15) for the potential of a dipole, that is,

\[
d\phi_A = \frac{P \ da \ dz \cos \theta}{4\pi \epsilon_0 r^2}.
\] (10.32)

The potential due to the entire column is

\[
\phi_A = \frac{P \ da}{4\pi \epsilon_0} \int_{z_1}^{z_2} \frac{dz \cos \theta}{r^2}.
\] (10.33)

This is simpler than it looks: \( dz \cos \theta \) is just \(-dr\), so that the integrand is a perfect differential, \( d(1/r) \). The result of the integration is then

\[
\phi_A = \frac{P \ da}{4\pi \epsilon_0} \left( \frac{1}{r_2} - \frac{1}{r_1} \right).
\] (10.34)

Equation (10.34) is precisely the same as the expression for the potential at \( A \) that would be produced by two point charges, a positive charge of magnitude \( P \ da \) sitting on top of the column at a distance \( r_2 \) from \( A \), and a negative charge of the same magnitude at the bottom of
10.7 The electric field caused by polarized matter

The electric field caused by polarized matter is equivalent, at least so far as its field at all external points is concerned, to two concentrated charges. Note that nowhere have we assumed that $A$ is far away from the column, that is, that $r_1$ and $r_2$ are much larger than the height of the column, $z_2 - z_1$. All that is required is that the distance from $A$ to any point in the column is much larger than the size of the dipoles (assumed to be very small) and also much larger than the width of the column (also assumed to be small), for then Eq. (10.32) will be valid.

We can prove Eq. (10.34) in another way without any mathematics. Consider a small section of the column of height $dz$, containing a dipole moment $P \, da \, dz$. Let us make an imitation or substitute for this by taking an unpolarized insulator of the same size and shape and sticking a charge $P \, da$ on top of it and a charge $-P \, da$ on the bottom. This little block now has the same dipole moment as that bit of our original column, and therefore it will make an identical contribution to the field at any remote point $A$. (The field inside our substitute, or very close to it, may be different from the field of the original – we don’t care about that.) Now make a whole set of such blocks and stack them up to imitate the polarized column; see Fig. 10.15(b). They must give the same field at $A$ as the whole column does, for each block gave the same contribution as its counterpart in the original. Now see what we have! At every joint the positive charge on the top of one block coincides with the negative charge on the bottom of the block above it, making the charge equal zero. The only charges left uncompensated are the negative charge $-P \, da$ on the bottom of the bottom block and the positive charge $P \, da$ on the top of the top block. Seen from a distant point such as $A$ (“distant” compared with the size of a block, not necessarily the whole column), these look like point charges. We conclude, as before, that two such charges produce at $A$ exactly the same field as does our whole column of polarized material.

With no further calculation we can extend this to a slab, or right cylinder, of any proportions uniformly polarized in a direction perpendicular to its parallel faces; see Fig. 10.16(a). The slab can simply be subdivided into a bundle of columns, and the potential outside will be the sum of the contributions of the columns, each of which can be replaced
Electric fields in matter

by a charge at either end. The charges on the top, \( P \, da \) on each column end of area \( da \), make up a uniform sheet of surface charge of density

\[
\sigma = P
\]

(10.35)

We conclude that the potential everywhere outside a uniformly polarized slab or cylinder (not necessarily far away) is precisely what would result from two sheets of surface charge located where the top and bottom surfaces of the slab were located, carrying the constant surface charge density \( \sigma = P \) and \( \sigma = -P \), respectively; see Fig. 10.16(b).

We are not quite ready to say anything about the field inside the slab. However, we do know the potential at all points on the surface of the slab – top, bottom, or sides. Any two such points, \( A \) and \( B \), can be connected by a path running entirely through the external field, so that the line integral \( \int E \cdot ds \) is entirely determined by the external field. It must be the same as the integral along the path \( A'B' \) in Fig. 10.16(b). A point literally on the surface of the dielectric might be within range of the intense molecular fields, the near field of the molecule that we have left out of our account. Let’s agree to define the boundary of the dielectric as a surface far enough out from the outermost atomic nucleus – 10 or 20 angstroms would be margin enough – so that at any point outside this boundary, the near fields of the individual atoms make a negligible contribution to the whole line integral from \( A \) to \( B \).

With this in mind, let’s look at a rather thin, wide plate of polarized material, of thickness \( t \), shown in cross section in Fig. 10.17(a). Figure 10.17(b) shows, likewise in cross section, the equivalent sheets of charge. For the system of two charge sheets, we know the field, of course, in the space both outside and between the sheets. The field strength inside, well away from the edges, must be just \( \sigma/\varepsilon_0 \), pointing down, and the potential difference between points \( A' \) and \( B' \) is therefore \( \sigma t/\varepsilon_0 \). The same potential difference must exist between corresponding points \( A \) and \( B \) on our polarized slab, because the entire external field is the same in the two systems.

**10.7.2 The field inside matter**

We can now address the field inside polarized matter. Is the internal field the same in the two systems in Fig. 10.17? Certainly not, because the slab is full of positive nuclei and electrons, with fields on the order of \( 10^{11} \) volts per meter pointing in one direction here, another direction there. But one thing is the same: the line integral of the field, reckoned over any internal path from \( A \) to \( B \), must be just \( \phi_B - \phi_A \), which, as we have seen, is the same as \( \phi_B' - \phi_A' \), which is equal to \( \sigma t/\varepsilon_0 \), or \( Pt/\varepsilon_0 \). This must be so because the introduction of atomic charges, no matter what their distribution, cannot destroy the conservative property of the electric field, expressed in the statement that \( \int E \cdot ds \) is independent of path, or \( \text{curl } E = 0 \).
10.7 The electric field caused by polarized matter

We know that in Fig. 10.17(b) the potential difference between the top and bottom sheets is nearly constant, except near the edges, because the interior electric field is practically uniform. Therefore in the central area of our polarized plate the potential difference between top and bottom must likewise be constant. In this region the line integral \( \int_A^B \mathbf{E} \cdot ds \) taken from any point \( A \) on top of the slab to any point \( B \) on the bottom, by any path, must always yield the same value \( Pt/\varepsilon_0 \). Figure 10.18 is a “magnified view” of the central region of the slab, in which the polarized molecules have been made to look something like H\(_2\)O molecules all pointing the same way. We have not attempted to depict the very intense fields that exist between and inside the molecules. (The field ten angstroms away from a water molecule is on the order of a hundred megavolts per meter, as you can discover from Fig. 10.14 and Eq. (10.18).) You must imagine some rather complicated field configurations in the neighborhood of each molecule. Now, the \( \mathbf{E} \) in \( \int \mathbf{E} \cdot ds \) represents the total electric field at a given point in space, inside or outside a molecule; it includes these complicated and intense fields just mentioned. We have reached the remarkable conclusion that any path through this welter of charges and fields, whether it dodges molecules or penetrates them, must yield the same value for the path integral, namely the value we find in the system of Fig. 10.17(b), where the field is quite uniform and has the strength \( P/\varepsilon_0 \).

This tells us that the spatial average of the electric field within our polarized slab must be \( -P/\varepsilon_0 \). By the spatial average of a field \( \mathbf{E} \) over some volume \( V \), which we might denote by \( \langle \mathbf{E} \rangle_V \), we mean precisely this:

\[
\langle \mathbf{E} \rangle_V = \frac{1}{V} \int_V \mathbf{E} \, dv.
\]  

(10.36)

One way to sample impartially the field in many equal volumes \( dv \) into which \( V \) might be divided would be to measure the field along each line in a bundle of closely spaced parallel lines. We have just seen that the line integral of \( \mathbf{E} \) along any or all such paths is the same as if we were in a constant electric field of strength \( -P/\varepsilon_0 \). That is the justification for the conclusion that, within the polarized dielectric slab of Figs. 10.17 and
Electric fields in matter

10.18, the spatial average of the field due to all the charges that belong to the dielectric is

\[ \langle E \rangle = -\frac{P}{\varepsilon_0} \]  

(10.37)

This average field is a macroscopic quantity. The volume over which we take the average should be large enough to include very many molecules, otherwise the average will fluctuate from one such volume to the adjoining one. The average field \( \langle E \rangle \) defined by Eq. (10.36) is really the only kind of macroscopic electric field in the interior of a dielectric that we can talk about. It provides the only satisfactory answer, in the context of a macroscopic description of matter, to the question, What is the electric field inside a dielectric material?

We may call the \( E \) in the integrand on the right, in Eq. (10.36), the microscopic field. If we imagine that we could measure the field values we need for the path integral, we will be measuring electric fields in vacuum, in the presence, of course, of electric charge. We will need very tiny instruments, for we may be called on to measure the field at a particular point just inside one end of a certain molecule. Have we any right to talk in this way about taking the line integral of \( E \) along some path that skirts the southwest corner of a particular molecule and then tunnels through its neighbor? Yes. The justification is the massive evidence that the laws of electromagnetism work down to a scale of distances much smaller than atomic size. We can even describe an experiment that would serve to measure the average of the microscopic electric field along a path defined well within the limits of atomic dimensions. All we have to do is shoot an energetic charged particle, an alpha particle for example, through the material. From the net change in its momentum, the average electric field that acted on it, over its whole path, could be inferred.

Let us review the properties of the average, or macroscopic, field \( \langle E \rangle \) defined by Eq. (10.36). Its line integral \( \int_A^B \langle E \rangle \cdot ds \) between any two points \( A \) and \( B \) that are reasonably far apart is independent of the path. It follows that curl \( \langle E \rangle = 0 \) and that \( \langle E \rangle \) is the negative gradient of a potential \( \langle \phi \rangle \). This potential function \( \langle \phi \rangle \) is itself a smoothed-out average, in the sense of Eq. (10.36), of the microscopic potential \( \phi \). (The latter rises to several million volts in the interior of every atomic nucleus!) The surface integral of \( \langle E \rangle \), \( \int \langle E \rangle \cdot da \), over any surface that encloses a reasonably large volume, is equal to \( 1/\varepsilon_0 \) times the charge within that volume.\(^3\) That is to say, \( \langle E \rangle \) obeys Gauss’s law, a statement we can also make in differential

\(^3\) We state this without proof, postponing consideration of the relation of the surface integral of an average field to the average of surface integrals of the microscopic field to Chapter 11, where the question arises in Section 11.8 in connection with the magnetic field inside matter. (See Fig. 11.18.)
form: \( \text{div}(\mathbf{E}) = \langle \rho \rangle / \epsilon_0 \), with the understanding that \( \langle \rho \rangle \) too is a local average over a suitably macroscopic volume. In short, the spatial average quantities \( \langle E \rangle \), \( \langle \phi \rangle \), and \( \langle \rho \rangle \) are related to one another in the same way as are the microscopic electric field, potential, and charge density in vacuum.

From now on, when we speak of the electric field \( \mathbf{E} \) inside any piece of matter much larger than a molecule, we will mean an average, or macroscopic, field as defined by Eq. (10.36), even when the brackets \( \langle \cdot \rangle \) are omitted.

### 10.8 Another look at the capacitor

At the beginning of this chapter we explained in a qualitative way how the presence of a dielectric between the plates of a capacitor increases its capacitance. Now we are ready to analyze quantitatively the dielectric-filled capacitor. What we have just learned about the electric field inside matter is the key to the problem. We identified as the macroscopic field \( \mathbf{E} \), the spatial average of the microscopic field. The line integral of that macroscopic \( \mathbf{E} \) between any two points \( A \) and \( B \) is path-independent and equal to the potential difference. Looking back at Fig. 10.2(a) we observe that the field \( \mathbf{E} \) in the empty capacitor must have had the value \( \phi_{12}/s \). But the potential difference between the plates, \( \phi_{12} \), which was established by the battery, was exactly the same in the dielectric-filled capacitor in Fig. 10.2(b). Hence the field \( \mathbf{E} \) in the dielectric, understood now as the macroscopic field, must have had the same value too, for it extends and is uniform over the same distance \( s \). (The layers in the diagram are actually negligible in thickness compared with \( s \).)

The fact that the \( \mathbf{E} \) fields are the same implies that the total charge on and near the top plate in the dielectric-filled capacitor must be the same as it was in the empty capacitor, namely \( Q_0 \). To prove that, we need only invoke Gauss’s law for a suitable imaginary box enclosing the charge layers, as indicated in Fig. 10.19. Now, the charge is made up of two parts, the charge on the plate \( Q \) (which will flow off when the capacitor is discharged) and \( Q' \), the charge that belongs to the dielectric. The charge on the plate is given by \( Q = \kappa Q_0 \). That was our definition of \( \kappa \). Therefore, if \( Q + Q' = Q_0 \) as we have just concluded, we must have

\[
Q' = Q_0 - Q = Q_0(1 - \kappa) \tag{10.38}
\]

We can think of this system as the superposition of a vacuum capacitor and a polarized dielectric slab, Fig. 10.19(b) and (c). In the vacuum capacitor with charge \( \kappa Q_0 \), the electric field \( E'' \) would be \( \kappa \) times the field \( E \). In the isolated polarized dielectric slab the field \( E' \) is \( -P/\epsilon_0 \), as stated in Eq. (10.37). The superposition of these two objects creates the actual field \( E \). That is,

\[
E = E'' + E' = \kappa E - \frac{P}{\epsilon_0} \tag{10.39}
\]
Electric fields in matter

Figure 10.19. The dielectric-filled capacitor of Fig. 10.2(b). The field \( E \), which is the average, or macroscopic, field in the dielectric, is \( \phi_{12}/s \), equal to the field in the empty capacitor of Fig. 10.2(a). The charge inside the Gauss box must equal \( Q_0 \), the charge on the plate of the empty capacitor. The system can be regarded as the superposition of a vacuum capacitor (b) and a polarized dielectric (c).

which can be rearranged like this:

\[
\frac{P}{\varepsilon_0 E} = \kappa - 1
\]  

(10.40)

The ratio \( P/\varepsilon_0 E \) (which is dimensionless) is called the electric susceptibility of the dielectric material and is denoted by \( \chi_e \) (Greek chi):

\[
\chi_e \equiv \frac{P}{\varepsilon_0 E} \implies P = \chi_e \varepsilon_0 E
\]  

(10.41)

From Eq. (10.40) we have

\[
\chi_e = \kappa - 1 \implies \kappa = 1 + \chi_e
\]  

(10.42)

In most materials under ordinary circumstances, it is the field \( \mathbf{E} \) in the dielectric that causes \( \mathbf{P} \). The relation is quite linear. That is to say, the electric susceptibility \( \chi_e \) is a constant characteristic of the particular material and is not dependent on the strength of the electric field or the size or shape of the electrodes. We call such materials, in which \( \mathbf{P} \) is proportional to \( \mathbf{E} \), linear dielectrics. Cases are known, however,
10.8 Another look at the capacitor

usually involving materials composed of polar molecules, in which polarization can be literally frozen in. A block of ice polarized by an externally applied electric field and then cooled in liquid helium will retain its polarization indefinitely after the external field is removed, thus providing a real example of the hypothetical polarized slab in Fig. 10.18.

In addition to frozen-in polarization, there are two other cases where the $P \propto E$ relation doesn’t hold. First, we can have a nonisotropic crystal, that is, one in which the polarization responds differently to electric fields in different directions. Each component of $P$ is then a (usually linear) function of, in general, all three components of $E$. In other words, the $P$ and $E$ vectors are related by a full matrix instead of a simple constant of proportionality. So they need not point in the same direction; see the discussion in Footnote 3 in Chapter 4 dealing with the analogous case involving $J$ and $E$. However, we will assume that the dielectrics in this chapter are isotropic, unless otherwise stated.

Second, the $P \propto E$ relation doesn’t hold if, as mentioned above, the proportionality factor $\chi_e\epsilon_0$ depends on the strength $E$ of the electric field. In this case $\chi_e\epsilon_0$ is a function of $E$, which means that $P$ is a non-linear function of $E$. If you want, you can consider the linear-dielectric $P = \chi_e\epsilon_0 E$ relation to be the first term in the Taylor series of $P$ as a function of $E$. But the nice thing is that, in most materials, this first term is all we need. As with isotropy, we will assume that our dielectrics are linear, unless otherwise stated. Note that in using definite values of $\kappa$ (that is, ones that are independent of $E$) to describe the various materials in Table 10.1, we are already assuming (correctly) that they are linear dielectrics. Furthermore, we are assuming (again correctly) that the materials are homogeneous and isotropic (on a macroscopic scale); that is, they have the same properties at all points and in all directions.

 Strictly speaking, filling a vacuum capacitor with dielectric material increases its capacitance by the precise factor $\kappa$ characteristic of that material only if we fill all the surrounding space too, or at least all the space where there is any electric field. In the example we discussed, it was tacitly assumed that the plates were so large compared with their distance of separation that “edge effects,” including the small amount of charge that would be on the outside of the plates near the edge (see Fig. 3.14(b)), could be neglected. A quite general statement can be made about a system of conductors of any shape or arrangement that is entirely immersed in a homogeneous dielectric – for instance, in a large tank of oil. With any charges, $Q_1, Q_2$, etc., on the various conductors, the macroscopic electric field $E_{med}$ at any location in the medium is just $1/\kappa$ times the field $E_{vac}$ that would exist at that location with the same charges on the same conductors in vacuum (Fig. 10.20). This has important consequences in semiconductors. When silicon, for example, is doped with phosphorus to make an $n$-type semiconductor, the high dielectric constant of the silicon crystal (see Table 10.1) greatly reduces the electrical attraction between the outermost electron of the phosphorus atom and the

Figure 10.20.
For the same charges on the conductors, the presence of the dielectric medium reduces all electric field intensities (and hence all potential differences) by the factor $1/\kappa$. The charges $Q_1$, $Q_2$, and $Q_3$ are the charges that would actually flow off the conductors if we were to discharge the system.
Electric fields in matter

rest of the atom. This makes it easy for the electron to leave the residual \( P^- \) ion and join the conduction band, as in Fig. 4.11(a).

This brings us to a more general problem. What if the space in our system is partly filled with dielectric and partly empty, with electric fields in both parts? We’ll begin with a somewhat artificial but instructive example, a polarized solid sphere in otherwise empty space.

### 10.9 The field of a polarized sphere

The solid sphere in Fig. 10.21(a) is supposed to be uniformly polarized, as if it had been carved out of the substance of the slab in Fig. 10.16(a). What must the electric field be like, both inside and outside the sphere? We take \( P \) as usual to denote the density of polarization, constant in magnitude and direction throughout the volume of the sphere. The polarized material could be divided, like the slab in Fig. 10.16(a), into columns parallel to \( P \), and each of these replaced by a charge of magnitude \( P \times (\text{column cross section}) \) at top and bottom. Thus the field we seek is that of a surface charge distribution spread over a sphere with density \( \sigma = P \cos \theta \). The factor \( \cos \theta \) enters, as should be evident from the figure, because a column of cross section \( da \) intercepts on the sphere a patch of surface of area \( da / \cos \theta \). Figure 10.21(b) is a cross section through this shell of equivalent surface charge in which the density of charge has been indicated by the varying thickness of the black semicircle above (positive charge density) and the light semicircle below (negative charge density).

If it has not already occurred to you, this figure may suggest that we think of the polarization \( P \) as having arisen from the slight upward displacement of a ball filled uniformly with positive charge of volume density \( \rho \), relative to a ball of negative charge of density \(-\rho\). That would leave uncompensated positive charge poking out at the top and negative charge showing at the bottom, varying in amount precisely as \( \cos \theta \) over the whole boundary. In the interior, where the positive and negative charge densities still overlap, they would exactly cancel one another. Taking this view, we see a very easy way to calculate the field outside the shell of surface charge. Any spherical charge distribution, as we know, has an external field the same as if its entire charge were concentrated at the center. So the superposition of two spheres of total charge \( Q \) and \(-Q\), with their centers separated by a small displacement \( s \), will produce an external field the same as that of two point charges \( Q \) and \(-Q\), a distance \( s \) apart. This is just a dipole with dipole moment \( p_0 = Qs \).

A microscopic description of the polarized substance leads us to the same conclusion. In Fig. 10.22(a) the molecular dipoles actually responsible for the polarization \( P \) have been crudely represented as consisting individually of a pair of charges \( q \) and \(-q\), a distance \( s \) apart, to make

---

\(4\) This follows from the fact that the thickness of the “semicircle” at a given point is the radial component of the vertical vector representing the displacement \( s \) of the top sphere relative to the bottom sphere. You can quickly show that this radial component is \( s \cos \theta \).
a dipole moment \( p = qs \). With \( N \) of these per cubic meter, we have \( P = Np = Nqs \), and the total number of such dipoles in the sphere is \((4\pi/3)r_0^3N\). The positive charges, considered separately (Fig. 10.22(b)), are distributed throughout a sphere with total charge content \( Q = (4\pi/3)r_0^3Nq \), and the negative charges occupy a similar sphere with its center displaced (Fig. 10.22(c)). Clearly each of these charge distributions can be replaced by a point charge at its center, if we are concerned with the field well outside the distribution. “Well outside” means far enough away from the surface so that the actual graininess of the charge distribution doesn’t matter, and of course that is something we always have to ignore when we speak of the macroscopic fields.

So, for present purposes, the picture of overlapping spheres of uniform charge density and the description in terms of actual dipoles in a vacuum are equivalent,\(^5\) and show that the field outside the distribution is the same as that of a single dipole located at the center. The moment of this dipole \( p_0 \) is simply the total polarization in the sphere:

\[
p_0 = Qs = \frac{4\pi}{3} r_0^3 Nqs = \frac{4\pi}{3} r_0^3 P. \tag{10.43}
\]

The quantities \( Q \) and \( s \) have, separately, no significance and may now be dropped from the discussion.

The external field of the polarized sphere is that of a central dipole \( p_0 \), not only at a great distance from the sphere but also right down to the surface, macroscopically speaking. All we had to do to construct Fig. 10.23, a representation of the external field lines, was to block out a circular area from Fig. 10.6.

The internal field is a different matter. Let’s look at the electric potential, \( \phi(x, y, z) \). We know the potential at all points on the spherical

\(^5\) This may have been obvious enough, but we have labored the details in this one case to allay any suspicion that the “smooth-charge-ball” picture, which is so different from what we know the interior of a real substance to be like, might be leading us astray.

**Figure 10.22.**
A sphere of lined-up molecular dipoles (a) is equivalent to superposed, slightly displaced, spheres of positive (b) and negative (c) charges.

**Figure 10.23.**
The field outside a uniformly polarized sphere is exactly the same as that of a dipole located at the center of the sphere.
Electric fields in matter

boundary because we know the external field. It is just the dipole potential, \( p_0 \cos \theta /4\pi \varepsilon_0 r^2 \), which on the spherical boundary of radius \( r_0 \) becomes

\[
\phi = p_0 \frac{\cos \theta}{4\pi \varepsilon_0 r_0^2} = \frac{Pr_0 \cos \theta}{3\varepsilon_0},
\]

where we have used Eq. (10.43). Since \( r_0 \cos \theta = z \), we see that the potential of a point on the sphere depends only on its \( z \) coordinate:

\[
\phi = \frac{Pz}{3\varepsilon_0}.
\]

The problem of finding the internal field has boiled down to this: Eq. (10.45) gives the potential at every point on the boundary of the region, inside which \( \phi \) must satisfy Laplace’s equation. According to the uniqueness theorem we proved in Chapter 3, that suffices to determine \( \phi \) throughout the interior. If we can find a solution, it must be the solution. Now the function \( Cz \), where \( C \) is any constant, satisfies Laplace’s equation, so Eq. (10.45) has actually handed us the solution to the potential in the interior of the sphere. That is, \( \phi_{\text{in}} = Pz/3\varepsilon_0 \). The electric field associated with this potential is uniform and points in the \(-z\) direction:

\[
E_z = \frac{\partial \phi_{\text{in}}}{\partial z} = -\frac{\partial}{\partial z} \left( \frac{Pz}{3\varepsilon_0} \right) = -\frac{P}{3\varepsilon_0}.
\]

As the direction of \( \mathbf{P} \) was the only thing that distinguished the \( z \) axis, we can write our result in more general form:

\[
\mathbf{E}_{\text{in}} = -\frac{\mathbf{P}}{3\varepsilon_0}
\]

This is the macroscopic field \( \mathbf{E} \) in the polarized material.

Figure 10.24 shows both the internal and external fields. At the upper pole of the sphere, the strength of the upward-pointing external field is, from Eq. (10.17) or Eq. (10.18) for the field of a dipole,

\[
E_z = \frac{2p_0}{4\pi \varepsilon_0 r^3} = \frac{2(4\pi r_0^3 P/3)}{4\pi \varepsilon_0 r_0^3} = \frac{2P}{3\varepsilon_0}
\]

(outside, at top),

which is just twice the magnitude of the downward-pointing internal field.

This example illustrates the general rules for the behavior of the field components at the surface of a polarized medium. \( \mathbf{E} \) is discontinuous at the boundary of a polarized medium, exactly as it would be at a surface in vacuum that carried a surface charge density \( \sigma = P_\perp \). The symbol \( P_\perp \) stands for the component of \( \mathbf{P} \) normal to the surface outward (which in the present case is \( P_\perp = P \cos \theta \)). It follows that \( E_\perp \), the normal component of \( \mathbf{E} \), must change abruptly by an amount \( P_\perp /\varepsilon_0 \); whereas \( E_\parallel \), the component of \( \mathbf{E} \) parallel to the boundary, remains continuous, that is, has the same value on both sides of the boundary (Fig. 10.25). Indeed,
at the north pole of our sphere, the net change in $E_z$ is $2P/3\epsilon_0 - (-P/3\epsilon_0)$, or $P/\epsilon_0$.

**Example (Continuity of $E_\parallel$)** For our polarized sphere, let's check that the component of $E$ parallel to the surface is continuous from inside to outside everywhere on the sphere. From Eq. (10.47) the internal field has magnitude $P/3\epsilon_0$ and points downward, so $E_\parallel^{\text{in}}$ is obtained by simply tacking on a factor of $\sin \theta$. That is, $E_\parallel^{\text{in}} = P \sin \theta/3\epsilon_0$. The tangential component of the external dipole field is given by the $E_\theta$ in Eq. (10.18):

$$E_\parallel^{\text{out}} = \frac{p_0 \sin \theta}{4\pi \epsilon_0 r^3} = \frac{(4\pi r_0^3 P/3) \sin \theta}{4\pi \epsilon_0 r_0^3} = \frac{P \sin \theta}{3\epsilon_0},$$

which equals $E_\parallel^{\text{in}}$, as desired.

Note that, for $0 < \theta < \pi$, the $\sin \theta$ factor is positive, so $E_\parallel^{\text{in}}$ and $E_\parallel^{\text{out}}$ point in the positive $\hat{\theta}$ direction, that is, away from the north pole. Similarly, for $\pi < \theta < 2\pi$, $E_\parallel^{\text{in}}$ and $E_\parallel^{\text{out}}$ point in the negative $\hat{\theta}$ direction, which again is away from the north pole (because positive $\hat{\theta}$ is directed clockwise around the full circle). A quick glance at Fig. 10.24 shows that the field lines are consistent with these facts.

The task of Exercise 10.36 is to use the explicit forms of the internal and external fields to show that $E_\perp$ has a discontinuity of $P_\perp/\epsilon_0$ everywhere on the surface of the sphere.

None of these conclusions depends on how the polarization of the sphere was caused. Assuming any sphere is uniformly polarized, Fig. 10.24 shows its field. Onto this can be superposed any field from other sources, thus representing many possible systems. This will not affect the discontinuity in $E$ at the boundary of the polarized medium. The above rules therefore apply in any system, the discontinuity in $E$ being determined solely by the existing polarization.

## 10.10 A dielectric sphere in a uniform field

As an example, let us put a sphere of dielectric material characterized by a dielectric constant $\kappa$ into a homogeneous electric field $E_0$ like the field between the parallel plates of a vacuum capacitor, Fig. 10.26. Let the sources of this field, the charges on the plates, be far from the sphere so that they do not shift as the sphere is introduced. Then whatever the field may be in the vicinity of the sphere, it will remain practically $E_0$ at a great distance. This is what is meant by putting a sphere into a uniform field. The total field $E$ is no longer uniform in the neighborhood of the sphere. It is the sum of the uniform field $E_0$ of the distant sources and a field $E'$ generated by the polarized matter itself:

$$E = E_0 + E'.$$  \hspace{1cm} (10.50)
Electric fields in matter

This relation is valid both inside and outside the sphere. The field \( E' \) depends on the polarization \( P \) of the dielectric, which in turn depends on the value of \( E \) inside the sphere:

\[
P = \chi_e \varepsilon_0 E_{\text{in}} = (\kappa - 1) \varepsilon_0 E_{\text{in}}. \tag{10.51}
\]

Remember that the \( E \) that appears in this expression involving \( \chi_e \) is the total electric field.

We don’t know yet what the total field \( E \) is; we know only that Eq. (10.51) has to hold at any point inside the sphere. If the sphere becomes uniformly polarized, an assumption that will need to be justified by our results, the relation between the polarization \( P \) of the sphere and its own field at points inside, \( E'_{\text{in}} \), is given by Eq. (10.47):\(^6\)

\[
E'_{\text{in}} = -\frac{P}{3\varepsilon_0}. \tag{10.52}
\]

Substituting the \( P \) from Eq. (10.51) into Eq. (10.52) quickly gives \( E'_{\text{in}} \) in terms of \( E_{\text{in}} \): we obtain \( E'_{\text{in}} = -(\kappa - 1)E_{\text{in}}/3 \). Substituting this into Eq. (10.50) gives the total field inside the sphere as

\[
E_{\text{in}} = E_0 - \frac{\kappa - 1}{3} E_{\text{in}} \implies E_{\text{in}} = \left(\frac{3}{2 + \kappa}\right) E_0 \tag{10.53}
\]

Because \( \kappa \) is greater than 1, the factor \( 3/(2 + \kappa) \) will be less than 1; the field inside the dielectric is weaker than \( E_0 \). The polarization is

\[
P = (\kappa - 1) \varepsilon_0 E_{\text{in}} \implies P = 3 \left(\frac{\kappa - 1}{\kappa + 2}\right) \varepsilon_0 E_0 \tag{10.54}
\]

The assumption of uniform polarization is now seen to be self-consistent.\(^7\)

To compute the total field \( E_{\text{out}} \) outside the sphere we must add vectorially to \( E_0 \) the field of a central dipole with dipole moment equal to \( P \) times the volume of the sphere. Some field lines of \( E \), both inside and outside the dielectric sphere, are shown in Fig. 10.27.

To summarize, we found \( E_{\text{in}} \) by effectively equating two different expressions for the field \( E'_{\text{in}} \) caused by the polarized matter. One expression is simply the statement of superposition, \( E'_{\text{in}} = E_{\text{in}} - E_0 \). The other expression is \( E'_{\text{in}} = -(\kappa - 1)E_{\text{in}}/3 \), which comes from the facts that \( E'_{\text{in}} \) is proportional to \( P \) (in the case of a sphere) and that \( P \) is proportional to \( E_{\text{in}} \) (in a linear dielectric).

\(^6\) In Eq. (10.47) we were using the symbol \( E_{\text{in}} \), without the prime, for this field. In that case it was the only field present.

\(^7\) That is what makes this system easy to deal with. For a dielectric cylinder of finite length in a uniform electric field, the assumption would not work. The field \( E' \) of a uniformly polarized cylinder – for instance one with its length about equal to its diameter – is not uniform inside the cylinder. (What must it look like?) Therefore \( E_{\text{in}} = E_0 + E'_{\text{in}} \) cannot be uniform – but in that case \( P = \chi_e E_{\text{in}} \) could not be uniform after all. In fact, it is only dielectrics of ellipsoidal shape, of which the sphere is a special case, that acquire uniform polarization in a uniform field.
10.11 The field of a charge in a dielectric medium, and Gauss’s law

Suppose that a very large volume of homogeneous linear dielectric has somewhere within it a concentrated charge $Q$, not part of the regular molecular structure of the dielectric. Imagine, for instance, that a small metal sphere has been charged and then dropped into a tank of oil. As was stated at the end of Section 10.8, the electric field in the oil is simply $1/\kappa$ times the field that $Q$ would produce in a vacuum:

$$E = \frac{Q}{4\pi\epsilon_0\kappa r^2}. \quad (10.55)$$

The product $\epsilon_0\kappa$ is commonly denoted by $\epsilon$, so we can write

$$E = \frac{Q}{4\pi\epsilon r^2} \quad \text{where} \quad \epsilon \equiv \kappa\epsilon_0 \quad \implies \quad \kappa = \frac{\epsilon}{\epsilon_0}. \quad (10.56)$$

The quantity $\epsilon$ is known as the permittivity of the dielectric. The vacuum permittivity, also called the permittivity of free space, is simply $\epsilon_0$.

It is interesting to see how Gauss’s law works out. The surface integral of $\mathbf{E}$ (which is the macroscopic, or space average, field, remember) taken over a sphere surrounding $Q$, gives $Q/\kappa\epsilon_0$, or $Q/\epsilon$, if we believe Eq. (10.55), and not $Q/\epsilon_0$. Why not? The answer is that $Q$ is not the only charge inside the sphere. There are also all the charges that make up the atoms and molecules of the dielectric. Ordinarily any volume of the oil would be electrically neutral. But now the oil is radially polarized, which means that the charge $Q$, assuming it is positive, has pulled in toward itself the negative charge in the oil molecules and pushed away the positive charges. Although the displacement may be only very slight in each molecule, still on the average any sphere we draw around $Q$ will contain more oil-molecule negative charge than oil-molecule positive charge. Hence the net charge in the sphere, including the “foreign” charge $Q$ at the center, is less than $Q$. In fact, it is $Q/\kappa$.

It is often useful to distinguish between the foreign charge $Q$ and the charges that make up the dielectric itself. Over the former we have some degree of control – charge can be added to or removed from an object, such as the plate of a capacitor. This is often called free charge. The other charges, which are integral parts of the atoms or molecules of the dielectric, are usually called bound charge. Structural charge might be a better name. These charges are not mobile; they are more or less elastically bound, contributing, by their slight displacement, to the polarization.

One can devise a vector quantity that is related by something like Gauss’s law to the free charge only. In the system we have just examined (a point charge $Q$ immersed in a dielectric), the vector $\kappa\mathbf{E}$ has this property. That is, $\int \kappa \mathbf{E} \cdot d\mathbf{a}$, taken over some closed surface $S$, equals $Q/\epsilon_0$ if $S$ encloses $Q$, and zero if it does not. By superposition, this must
Electric fields in matter

hold for any collection of free charges described by a free-charge density \( \rho_{\text{free}}(x, y, z) \) in an infinite homogeneous linear dielectric medium:

\[
\int_S \kappa \mathbf{E} \cdot d\mathbf{a} = \frac{1}{\epsilon_0} \int_V \rho_{\text{free}} \, dv, \tag{10.57}
\]

where \( V \) is the volume enclosed by the surface \( S \). An integral relation like this implies a “local” relation between the divergence of the vector field \( \kappa \mathbf{E} \) and the free charge density:

\[
\text{div} (\kappa \mathbf{E}) = \frac{\rho_{\text{free}}}{\epsilon_0}. \tag{10.58}
\]

Since \( \kappa \) has been assumed to be constant throughout the medium, Eq. (10.58) tells us nothing new. However, it can help us to isolate the role of the bound charge. In any system whatsoever, the fundamental relation (namely Gauss’s law) between electric field \( \mathbf{E} \) and total charge density \( \rho_{\text{free}} + \rho_{\text{bound}} \) remains valid:

\[
\text{div} \mathbf{E} = \frac{1}{\epsilon_0} (\rho_{\text{free}} + \rho_{\text{bound}}). \tag{10.59}
\]

Subtracting Eq. (10.59) from Eq. (10.58) yields

\[
\text{div} (\kappa - 1) \mathbf{E} = -\frac{\rho_{\text{bound}}}{\epsilon_0}. \tag{10.60}
\]

According to Eq. (10.40), \((\kappa - 1) \mathbf{E} = \mathbf{P}/\epsilon_0\) for a linear dielectric, so Eq. (10.60) implies that

\[
\text{div} \mathbf{P} = -\rho_{\text{bound}} \tag{10.61}
\]

Equation (10.61) states a local relation. It cannot depend on conditions elsewhere in the system, nor on how the particular arrangement of bound charges is maintained. Any arrangement of bound charge that has a certain local excess, per unit volume, of nuclear protons over atomic electrons must represent a polarization with a certain divergence. So, although we derived Eq. (10.61) by using relations pertaining to linear dielectrics, it must in fact hold universally, not just in an unbounded linear dielectric. It doesn’t matter how the polarization comes about. (See Problem 10.11 for a general proof.) You can get a feeling for the identity expressed in Eq. (10.61) by imagining a few polar molecules arranged to give a polarization with a positive divergence (Fig. 10.28). The dipoles point outward, which necessarily leaves a little concentration of negative charge in the middle. Of course, Eq. (10.61) refers to averages over volume elements so large that \( \mathbf{P} \) and \( \rho_{\text{bound}} \) can be treated as smoothly varying quantities.

From Eqs. (10.59) and (10.61), both of which are true in any system whatsoever, we get the relation

\[
\text{div} (\epsilon_0 \mathbf{E} + \mathbf{P}) = \rho_{\text{free}}. \tag{10.62}
\]
This is quite independent of any relation between $E$ and $P$; it is not limited to linear dielectrics (where $P$ is proportional to $E$).

It is customary to give the combination $\varepsilon_0 E + P$ a special name, the electric displacement vector, and its own symbol, $D$. That is, we define $D$ by

$$D \equiv \varepsilon_0 E + P$$

and Eq. (10.62) becomes

$$\text{div} D = \rho_{\text{free}}$$

This relation, or equivalently Eq. (10.62), holds in any situation in which the macroscopic quantities $P$, $E$, and $\rho$ can be defined.

If additionally we are dealing with a linear dielectric, then by comparing Eqs. (10.58) and (10.64) we see that $D$ is simply $\kappa \varepsilon_0 E$, or

$$D = \varepsilon E$$

(for a linear dielectric).

This alternatively follows from Eq. (10.63) by using Eq. (10.41) to write $P$ as $\chi \varepsilon_0 E$, and then using Eq. (10.42) to write $1 + \chi \varepsilon_0$ as $\kappa$.

The appearance of Eq. (10.64) may suggest that we should look on $D$ as a vector field whose source is the free charge distribution $\rho_{\text{free}}$ (up to a factor of $\varepsilon_0$), in the same sense that the total charge distribution $\rho$ is the source of $E$. That would be wrong. The electrostatic field $E$ is uniquely determined – except for the addition of a constant field – by the charge distribution $\rho$ because, supplementing the law $\text{div} E = \rho/\varepsilon_0$, there is another universal condition, $\text{curl} E = 0$. It is not true, in general, that $\text{curl} D = 0$. Thus the distribution of free charge is not sufficient to determine $D$ through Eq. (10.64). Something else is needed, such as the boundary conditions at various dielectric surfaces. The boundary conditions on $D$ are of course merely an alternative way of expressing the boundary conditions involving $E$ and $P$, already stated near the end of Section 10.9 and in Fig. 10.25.

**Example (Continuity of $D_\perp$)** For our polarized sphere in Section 10.9, we saw that $E_\parallel$ was continuous across the boundary whereas $E_\perp$ was not. These boundary conditions hold for any shape of polarized material. It turns out that the opposite conditions are true for $D$. That is, $D_\perp$ is continuous across the boundary whereas $D_\parallel$ is not. You can derive these boundary conditions in Problem 10.12. For now, let’s just verify that $D_\perp$ is continuous across the boundary of our polarized sphere.

Inside the sphere, we have $E = -P/3\varepsilon_0$, so the displacement vector is $D = \varepsilon_0 (-P/3\varepsilon_0) + P = 2P/3$. The radial component of this is

$$D^\text{in}_\perp = D^\text{in}_r = \frac{2P \cos \theta}{3}.$$ 

Outside the sphere, $E$ is the field due to a dipole with $p_0 = (4\pi R^3/3)P$. The radial component of the dipole field is $E_r = p_0 \cos \theta / 2\pi \varepsilon_0 R^3$. In terms of $P$ this
becomes $E_r = 2P \cos \theta / 3\epsilon_0$. Since $P = 0$ outside the sphere, the external $D$ is obtained by simply multiplying the external $E$ by $\epsilon_0$. Therefore

$$D_{\text{out}}^\perp \equiv D_r^\text{out} = \frac{2P \cos \theta}{3}.$$  

(10.67)

This equals the above $D_{\perp}^{\text{in}}$, as desired.

The task of Exercise 10.41 is to use the explicit forms of the internal and external fields to find the discontinuity in $D_{\parallel}$ everywhere on the surface of the sphere.

In the approach we have taken to electric fields in matter, the introduction of $D$ is an artifice that is not, on the whole, very helpful. We have mentioned $D$ because it is hallowed by tradition, beginning with Maxwell,\(^8\) and the student is sure to encounter it in other books, many of which treat it with more respect than it deserves.

Our essential conclusions about electric fields in matter can be summarized as follows:

1. Matter can be polarized, its condition being described completely, so far as the macroscopic field is concerned, by a polarization density $P$, which is the dipole moment per unit volume. The contribution of such matter to the electric field $E$ is the same as that of a charge distribution $\rho_{\text{bound}}$, existing in vacuum and having the density $\rho_{\text{bound}} = -\text{div} P$. In particular, at the surface of a polarized substance, where there is a discontinuity in $P$, this reduces to a surface charge of density $\sigma = -\Delta P_{\perp}$. Add any free charge distribution that may be present, and the electric field is the field that this total charge distribution would produce in vacuum. This is the macroscopic field $E$ both inside and outside matter, with the understanding that inside matter it is the spatial average of the true microscopic field.

2. If $P$ is proportional to $E$ in a material, we call the material a linear dielectric. We define the electric susceptibility $\chi_e$ and the dielectric constant $\kappa$ characteristic of that material as $\chi_e = P/\epsilon_0 E$ and $\kappa = 1 + \chi_e$. Free charges immersed in a linear dielectric give rise to electric fields that are $1/\kappa$ times as strong as the same charges would produce in vacuum.

**10.12 A microscopic view of the dielectric**

The polarization $P$ in the dielectric is simply the large-scale manifestation of the electric dipole moments of the atoms or molecules of which

\(^8\) The prominence of $D$ in Maxwell’s formulation of electromagnetic theory, and his choice of the name *displacement*, can perhaps be traced to his inclination toward a kind of mechanical model of the “aether.” Whittaker has pointed out in his classic text (Whittaker, 1960) that this inclination may have led Maxwell himself astray at one point in the application of his theory to the problem of reflection of light from a dielectric.
the material is composed. \( \mathbf{P} \) is the mean dipole moment density, the total vector dipole moment per unit volume – averaged, of course, over a region large enough to contain an enormous number of atoms. If there is no electric field to establish a preferred direction, \( \mathbf{P} \) will be zero. That will surely be true for an ordinary liquid or a gas, and for solids too if we ignore the possibility of “frozen-in” polarization mentioned in Section 10.8. In the presence of an electric field in the medium, polarization can arise in two ways. (1) Every atom or molecule will acquire an induced dipole moment proportional to, and in the direction of, the field \( \mathbf{E} \) that acts on that atom or molecule. (2) If molecules with permanent dipole moments are present in the medium, their orientations will no longer be perfectly random; alignment of their dipole moments in the field direction will be favored slightly over alignment in the opposite direction. Both effects (1) and (2) lead to polarization in the direction \( \mathbf{E} \), that is, to a positive value of \( \chi_e \equiv \mathbf{P}/\epsilon_0 \mathbf{E} \), the electric susceptibility.

Let us consider first the induced atomic moments in a medium in which the atoms or molecules are rather far apart. An example is a gas at atmospheric density, in which there are something like \( 3 \cdot 10^{25} \) molecules per cubic meter. We shall assume that the field \( \mathbf{E} \) that acts on an individual molecule is the same as the average, or macroscopic, field \( \mathbf{E} \) in the medium. In making this assumption, we are neglecting the field at a molecule that is produced by the induced dipole moment of a nearby molecule. Let \( \alpha \) be the polarizability of every molecule and \( N \) the mean number of molecules per cubic meter. The dipole moment induced in each molecule is \( \mathbf{p} = \alpha \mathbf{E} \), and the resulting polarization of the medium, \( \mathbf{P} \), is simply

\[
\mathbf{P} = N \mathbf{p} = N \alpha \mathbf{E}.
\]  

(10.68)

This gives us at once the electric susceptibility \( \chi_e \),

\[
\chi_e = \frac{\mathbf{P}}{\epsilon_0 \mathbf{E}} = \frac{N \alpha}{\epsilon_0},
\]  

(10.69)

and the dielectric constant \( \kappa \),

\[
\kappa = 1 + \chi_e = 1 + \frac{N \alpha}{\epsilon_0}.
\]  

(10.70)

The methane molecule in Fig. 10.12 has a polarizability value (or rather an \( \alpha/4\pi \epsilon_0 \) value) of \( 2.6 \cdot 10^{-30} \) m\(^3\). At standard conditions of 0 °C and atmospheric pressure there are approximately \( 2.8 \cdot 10^{25} \) molecules in 1 m\(^3\). According to Eq. (10.70), the dielectric constant of methane at that density ought to have the value

\[
\kappa = 1 + \frac{N \alpha}{\epsilon_0} = 1 + \frac{1}{\epsilon_0} (2.8 \cdot 10^{25} \text{ m}^{-3}) (4\pi \epsilon_0 \cdot 2.6 \cdot 10^{-30} \text{ m}^3)
\]

\[
= 1.00091.
\]  

(10.71)
Up to rounding errors in the numbers we used, this agrees with the value of $\kappa$ listed for methane in Table 10.1. The agreement is hardly surprising, for the value of $\alpha/4\pi \varepsilon_0$ given in Fig. 10.12 was probably deduced originally by applying the simple theory we have just developed to an experimentally measured dielectric constant.

We have already noted in Section 10.5 that the atomic polarizability $\alpha/4\pi \varepsilon_0$, which has the dimensions of volume, is in order of magnitude about equal to the volume of an atom. That being so, the product $N\alpha/4\pi \varepsilon_0$, which is just $\chi_e/4\pi$ according to Eq. (10.69), is about equal to the fraction of the volume of the medium that is taken up by atoms. Now the density of a gas under standard conditions is roughly one-thousandth of the density of the same substance condensed to liquid or solid. In the case of methane the ratio is close to $1/1000$; in the case of air, $1/700$. The gas is about 99.9 percent empty space. In the solid or liquid, on the other hand, the molecules are practically touching one another. The fraction of the volume they occupy is not much less than unity. This tells us that, in condensed matter generally, the induced polarization will result in a value of $\chi_e/4\pi$ of order of magnitude unity. In fact, as our brief list in Table 10.1 suggests, and as a more extensive tabulation would confirm, the value of $\chi_e/4\pi = (\kappa - 1)/4\pi$ for most nonpolar liquids and solids ranges from about 0.1 to 1. We can now see why.

We can see, too, why an exact theory of the susceptibility $\chi_e$ of a solid or liquid is not so easy to develop. When the atoms are crowded together until they almost “touch,” the effect of one atom on its neighbors cannot be neglected. The distance $b$ between nearest neighbors is approximately $N^{-1/3}$. Let an electric field $E$ induce a dipole moment $p = E\alpha$ in each atom. This dipole $p$ on one atom will cause a field of strength $E' \sim p/4\pi \varepsilon_0 b^3$ at the location of the next atom. But $1/b^3 \approx N$, hence $E' \sim E\alpha N/4\pi \varepsilon_0$. As we have just explained, in condensed matter $\alpha N/4\pi \varepsilon_0$ is necessarily of order unity. Hence $E'$ is not small, and certainly not negligible, compared with $E$. Just what the effective field is that polarizes an atom in this situation is a question with no very obvious answer.9

Molecules with permanent electric dipole moments, polar molecules, respond to an electric field by trying to line up parallel to it. So long as the dipole moment $p$ is not pointing in the direction of $E$, there is a torque $p \times E$ tending to turn $p$ into the direction of $E$. (Look back at Eq. (10.21) and Fig. 10.8(b).) Of course, the torque is zero if $p$ happens to be pointing exactly opposite to $E$, but that condition is unstable. Torque on the electric dipole is torque on the molecule itself. A state of lowest energy will have been attained if and when all the polar molecules have rotated to bring their dipole moments into the $E$ direction. While settling down to that state of perfect alignment they will have given off some energy,

9 An elementary, approximate, treatment of this problem, leading to what is called the Clausius–Mossotti relation, can be found in Section 9.13 in the first edition of this book.
through rotational friction, to their surroundings. The resulting polarization would be gigantic. In water there are about \(3.3 \cdot 10^{28}\) molecules/m\(^3\); the dipole moment of each (Fig. 10.14) is \(6.1 \cdot 10^{-30}\) C-m. With complete alignment of the dipoles, \(P\) would be 0.2 C/m\(^2\). If Fig. 10.24 were a picture of a water droplet thus polarized, the field strength just outside the drop, of order \(P/\epsilon_0\) from Eq. (10.48), would exceed \(10^{10}\) V/m!

This does not happen. Nothing approaching complete alignment is attained in any reasonable applied field \(E\). Why not? The reason is essentially the same as the reason why the molecules of air in a room are not found all lying on the floor – which is, after all, the arrangement of lowest potential energy. We must think about temperature and about the energy of thermal agitation that every molecule exhibits at a given absolute temperature \(T\). In magnitude, that energy is \(kT\), where \(k\) is the universal constant called Boltzmann's constant. At room temperature, \(kT\) amounts to \(4 \cdot 10^{-21}\) joule. In a system all at temperature \(T\), the mean translational energy of a molecule – or, for that matter, of any object small or large – is \((3/2)kT\). More to the point here, the mean rotational energy of a molecule is just \(kT\). Now, the air molecules do not all gather near the floor because the change in gravitational potential energy in elevating by a couple of meters a molecule of mass \(5 \cdot 10^{-26}\) kg is only, as you can readily compute, about \(10^{-24}\) joule, less than \(1/1000\) of \(kT\). On the other hand, the air near the floor is slightly more dense than the air near the ceiling, even when there is no temperature gradient. That is just the well-known change of barometric pressure with height. Air near the floor is fractionally more dense (when the difference is slight) by just \(mgh/kT\), \(mgh\) being the difference in gravitational potential energy between the two levels.

Similarly, in our dielectric we shall find a slight excess of molecular dipoles in the orientation of lower potential energy, that is, pointing in the direction of \(E\), or with a component in that direction. The fractional excess in the favored directions will be, in order of magnitude, \(pE/kT\). The numerator represents the difference in potential energy. Actually the work required to turn a dipole from the direction of \(E\) to the opposite direction is \(2pE\) (see Eq. (10.22)), but averaging over angles would bring in other numerical factors that we are leaving out. With \(N\) dipoles per unit volume, the polarization \(P\), which would be \(Np\) if they were totally aligned, will be smaller by something like the factor \(pE/kT\). The polarization to be expected is therefore, in order of magnitude,

\[
P \approx Np \left(\frac{pE}{kT}\right) = \frac{Np^2}{kT} E, \quad (10.72)
\]

and the susceptibility is

\[
\chi_e = \frac{P}{\epsilon_0 E} \approx \frac{Np^2}{\epsilon_0 kT}. \quad (10.73)
\]
For water at room temperature, the quantity on the right in Eq. (10.73) is about 35, whereas with $\kappa = 80$, the actual value of $\chi_e$ is 79. Evidently a factor of roughly 2.3 is needed on the right in Eq. (10.73), in this case, to convert our order-of-magnitude estimate into a correct prediction. Deriving that factor theoretically is quite difficult, for the interactions of neighboring molecules complicate matters even more than in the case of the nonpolar dielectric.

If you apply an electric field of $10^4$ V/m to water, the resulting polarization, $P = \chi_e \varepsilon_0 E = 7 \cdot 10^{-6}$ C/m$^2$, is equivalent to the alignment of $1.1 \cdot 10^{24}$ H$_2$O dipoles per cubic meter, or about one molecule in 30,000. Even so, this polarization is an order of magnitude greater than that caused by the same field in any nonpolar dielectric.

10.13 Polarization in changing fields

So far we have considered only electrostatic fields in matter. We need to look at the effects of electric fields that are varying in time, like the field in a capacitor used in an alternating-current circuit. The important question is, will the changes in polarization keep up with the changes in the field? Will the ratio of $P$ to $E$, at any instant, be the same as in a static electric field? For very slow changes we should expect no difference but, as always, the criterion for slowness depends on the particular physical process. It turns out that induced polarization and the orientation of permanent dipoles are two processes with quite different response times.

The induced polarization of atoms and molecules occurs by the distortion of the electronic structure. Little mass is involved, and the structure is very stiff; its natural frequencies of vibration are extremely high. To put it another way, the motions of the electrons in atoms and molecules are characterized by periods on the order of $10^{-16}$ s—something like the period of a visible light wave. To an atom, $10^{-14}$ s is a long time. It has no trouble readjusting its electronic structure in a time like that. Because of this, strictly nonpolar substances behave practically the same from direct current (zero frequency) up to frequencies close to those of visible light. The polarization keeps in step with the field, and the susceptibility $\chi_e = P/\varepsilon_0 E$ is independent of frequency.

The orientation of a polar molecule is a process quite different from the mere distortion of the electron cloud. The whole molecular framework has to rotate. On a microscopic scale, it is rather like turning a peanut end for end in a bag of peanuts. The frictional drag tends to make the rotation lag behind the torque and to reduce the amplitude of the resulting polarization. Where on the time scale this effect sets in varies enormously from one polar substance to another. In water, the “response time” for dipole reorientation is something like $10^{-11}$ s. The dielectric constant remains around 80 up to frequencies on the order of $10^{10}$ Hz. Above $10^{11}$ Hz, $\kappa$ falls to a modest value typical of a nonpolar liquid. The dipoles simply cannot follow so rapid an alternation of the field.
In other substances, especially solids, the characteristic time can be much longer. In ice just below the freezing point, the response time for electrical polarization is around $10^{-5}$ s. Figure 10.29 shows some experimental curves of dielectric constant versus frequency for water and ice.

10.14 The bound-charge current

Wherever the polarization in matter changes with time, there is an electric current, a genuine motion of charge. Suppose there are $N$ dipoles per cubic meter of dielectric, and that in the time interval $dt$ each changes from $p$ to $p + dp$. Then the macroscopic polarization density $P$ changes from $P = Np$ to $P + dP = N(p + dp)$. Suppose the change $dp$ was effected by moving a charge $q$ through a distance $ds$, in each atom: $q ds = dp$. Then during the time $dt$ there was actually a charge cloud of density $\rho = Nq$, moving with velocity $v = ds/dt$. This is a conduction current of a certain density $J$ in coulombs per second per square meter:

$$J = \rho v = Nq \frac{ds}{dt} = N \frac{dp}{dt} \implies J = \frac{dP}{dt} \quad (10.74)$$

The connection between rate of change of polarization and current density, $J = dP/dt$, is independent of the details of the model. A changing polarization is a conduction current, not essentially different from any other. Note that if we take the divergence of both sides of Eq. (10.74) and use Eq. (10.61), we obtain $\text{div} \ J = d(\text{div} \ P)/dt = -d\rho_{\text{bound}}/dt$, which is consistent with the continuity equation in Eq. (4.10).

Naturally, such a current is a source of magnetic field. If there are no other currents around, we should write Maxwell’s equation, $\text{curl} \ B = \mu_0 \varepsilon_0 (\partial E/\partial t) + \mu_0 J$, as

$$\text{curl} \ B = \mu_0 \varepsilon_0 \frac{\partial E}{\partial t} + \mu_0 \frac{\partial P}{\partial t}. \quad (10.75)$$
The only difference between an “ordinary” conduction current density and the current density $\partial P/\partial t$ is that one involves free charge in motion, the other bound charge in motion. There is one rather obvious practical distinction – you can’t have a steady bound-charge current, one that goes on forever unchanged. Usually we prefer to keep account separately of the bound-charge current and the free-charge current, retaining $J$ as the symbol for the free-charge current density only. Then to include all the currents in Maxwell’s equation we have to write it this way:

$$\text{curl } B = \mu_0 \epsilon_0 \frac{\partial E}{\partial t} + \mu_0 \frac{\partial P}{\partial t} + \mu_0 J. \quad (10.76)$$

In a linear dielectric medium, we can write $\epsilon_0 E + P = \epsilon E$, allowing a shorter version of Eq. (10.76):

$$\text{curl } B = \mu_0 \epsilon \frac{\partial E}{\partial t} + \mu_0 J. \quad (10.77)$$

More generally, Eq. (10.76) can also be abbreviated by introducing the vector $D$, previously defined in any medium as $\epsilon_0 E + P$ (which reduces to $\epsilon E$ in a linear dielectric):

$$\text{curl } B = \mu_0 \frac{\partial D}{\partial t} + \mu_0 J \quad (10.78)$$

The term $\partial D/\partial t$ is usually referred to as the displacement current. Actually, the part of it that involves $\partial P/\partial t$ represents, as we have seen, an honest conduction current, real charges in motion. The only part of the total current density that is not simply charge in motion is the $\partial E/\partial t$ part, the true vacuum displacement current which we discussed in Chapter 9. Incidentally, if we want to express all components of the full current density in units corresponding to those of $J$, we can factor out the $\mu_0$ and write Eq. (10.76) as follows:

$$\text{curl } B = \mu_0 \left( \epsilon_0 \frac{\partial E}{\partial t} + \frac{\partial P}{\partial t} + J \right). \quad (10.79)$$

Involved in the distinction between bound charge and free charge is a question we haven’t squarely faced: can one always identify unambiguously the “molecular dipole moments” in matter, especially solid matter? The answer is no. Let us take a microscopic view of a thin wafer of sodium chloride crystal. The arrangement of the positive sodium ions...
and the negative chlorine ions was shown in Fig. 1.7. Figure 10.30 is a cross section through the crystal, which extends on out to the right and the left. If we choose to, we may consider an adjacent pair of ions as a neutral molecule with a dipole moment. Grouping them as in Fig. 10.30(a), we describe the medium as having a uniform macroscopic polarization density $P$, a vector directed downward. At the same time, we observe that there is a layer of positive charge over the top of the crystal and a layer of negative charge over the bottom, which, not having been included in our molecules, must be accounted free charge.

Now we might just as well have chosen to group the ions as in Fig. 10.30(b). According to that description, $P$ is a vector upward, but we have a negative free-charge layer on top of the crystal and a positive free-charge layer beneath. Either description is correct. You will have no trouble finding another one, also correct, in which $P$ is zero and there is no free charge. Each description predicts $E = 0$. The macroscopic field $E$ is an observable physical quantity. It can depend only on the charge distribution, not on how we choose to describe the charge distribution.

This example teaches us that in the real atomic world the distinction between bound charge and free charge is more or less arbitrary, and so, therefore, is the concept of polarization density $P$. The molecular dipole is a well-defined notion only where molecules as such are identifiable – where there is some physical reason for saying, “This atom belongs to this molecule and not to that.” In many crystals such an assignment is meaningless. An atom or ion may interact about equally strongly with all its neighbors; one can only speak of the whole crystal as a single molecule.

10.15 An electromagnetic wave in a dielectric

In Eq. (9.17) we wrote out Maxwell’s equations for the electric and magnetic fields in vacuum, including source terms – charge density $\rho$ and current density $J$. Now we want to consider an electromagnetic field in an unbounded dielectric medium. The dielectric is a perfect insulator,
Electric fields in matter

we shall assume, so there is no free-charge current. That is, the last term on the right in Eqs. (10.76) through (10.79), the free-charge current density \( \mathbf{J} \), will be zero. No free charge is present either, but there could be a nonzero density of bound charge if \( \text{div} \mathbf{E} \) is not zero. Let us agree to consider only fields with \( \text{div} \mathbf{E} = 0 \). Then \( \rho \), both bound and free, will be zero throughout the medium. No change is called for in the first induction equation, \( \text{curl} \mathbf{E} = -\partial \mathbf{B}/\partial t \). For the second equation, we now take Eq. (10.77) without the free-charge-current term: \( \text{curl} \mathbf{B} = \mu_0 \varepsilon (\partial \mathbf{E}/\partial t) \). The dielectric constant \( \varepsilon \) takes account of the bound-charge current as well as the vacuum displacement current. Our complete set of equations has become

\[
\text{curl} \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \text{div} \mathbf{E} = 0; \\
\text{curl} \mathbf{B} = \mu_0 \varepsilon \frac{\partial \mathbf{E}}{\partial t}, \quad \text{div} \mathbf{B} = 0. \quad (10.80)
\]

These differ from Eq. (9.18) only in the replacement of \( \varepsilon_0 \) with \( \varepsilon \) in the second induction equation.

As we did in Section 9.4, let us construct a wavelike electromagnetic field that can be made to satisfy Maxwell’s equations. This time we give our trial wave function a slightly more general form:

\[
\mathbf{E} = \hat{z} E_0 \sin(ky - \omega t), \\
\mathbf{B} = \hat{x} B_0 \sin(ky - \omega t). \quad (10.81)
\]

The angle \( ky - \omega t \) is called the phase of the wave. For a point that moves in the positive \( y \) direction with speed \( \omega/k \), the phase \( ky - \omega t \) remains constant. In other words, \( \omega/k \) is the phase velocity of this wave. This term is used when it is necessary to distinguish between two velocities, phase velocity and group velocity. There is no difference in the case we are considering, so we shall call \( \omega/k \) simply the wave velocity, the same as \( v \) in our discussion in Section 9.4. At any fixed location, such as \( y = y_0 \), the fields oscillate in time with angular frequency \( \omega \). At any instant of time, such as \( t = t_0 \), the phase differs by \( 2\pi \) at planes one wavelength \( \lambda \) apart, where \( \lambda = 2\pi/k \).

The divergence equations in Eq. (10.80) are quickly seen to be satisfied by the wave in Eq. (10.81). For the curl equations, the space and time derivatives we need are those listed in Eq. (9.24) with small alterations:

\[
\text{curl} \mathbf{E} = \hat{z} E_0 k \cos(ky - \omega t), \quad \frac{\partial \mathbf{E}}{\partial t} = -\hat{z} E_0 \omega \cos(ky - \omega t), \\
\text{curl} \mathbf{B} = -\hat{x} B_0 k \cos(ky - \omega t), \quad \frac{\partial \mathbf{B}}{\partial t} = -\hat{x} B_0 \omega \cos(ky - \omega t). \quad (10.82)
\]
Substituting these into Eq. (10.80), we find that the curl equations are satisfied if

\[
\frac{\omega}{k} = \pm \frac{1}{\sqrt{\mu_0 \epsilon}} \quad \text{and} \quad E_0 = \pm \frac{B_0}{\sqrt{\mu_0 \epsilon}}
\] (10.83)

The wave velocity \( v = \omega/k \) differs from the velocity of light in vacuum (which is \( c = 1/\sqrt{\mu_0 \epsilon_0} \)) by the factor \( \sqrt{\epsilon_0/\epsilon} = 1/\sqrt{\kappa} \). Since \( \kappa > 1 \), we have \( v < c \). The electric and magnetic field amplitudes, \( E_0 \) and \( B_0 \), which were related by \( E_0 = cB_0 \) for the wave in vacuum, here are related by \( E_0 = vB_0 \), where \( v = 1/\sqrt{\mu_0 \epsilon} \). For a given magnetic amplitude \( B_0 \), the electric amplitude \( E_0 \) is smaller in the dielectric than in vacuum. In other respects the wave resembles our plane wave in vacuum: \( B \) is perpendicular to \( E \), and the wave travels in the direction of \( E \times B \). Of course, if we compare a wave in a dielectric with a wave of the same frequency in vacuum, the wavelength \( \lambda \) in the dielectric will be less than the vacuum wavelength by \( 1/\sqrt{\kappa} \) since \( \text{frequency} \times \text{wavelength} = \text{velocity} \).

Light traveling through glass provides an example of the wave just described. In optics it is customary to define \( n \), the index of refraction of a medium, as the ratio of the speed of light in vacuum to the speed of light in that medium. We have now discovered that \( n \) is nothing more than \( \sqrt{\kappa} \). In fact, we have now laid most of the foundation for a classical theory of optics. Of course, we must be careful to use the appropriate value of \( \kappa \). Take water, for example. If we use the \( \kappa = 80 \) value from Table 10.1, we obtain \( n \approx 9 \). But the actual index of refraction of water is \( n = 1.33 \). What’s going on here? \textit{Hint}: The answer is contained in a figure in this chapter.

\section{10.16 Applications}

The \textit{pollination} of flowers by bees is helped by polarization effects. When bees travel through the air, they become positively charged due to triboelectric effects with the air; air molecules strip off electrons from the bee when its wings collide with the molecules. When the bee gets close to the pollen on the anther of the flower, the bee’s charge polarizes the pollen, which then experiences a dipole attraction toward the bee. The pollen jumps to the bee and lands on the bee’s hairs (while maintaining zero net charge, because the hairs aren’t conductive). When the bee then gets close to the stigma of a flower, it induces a negative charge on the stigma. The electric field from the somewhat pointed stigma wins out over the field from the somewhat rounded bee, so the pollen jumps to the stigma. Note the lack of symmetry between the anther and the stigma; the anther gives up the pollen, while the stigma attracts it. This lack of symmetry arises mainly from the fact that the stigma has a more conductive path to ground than the anther has, so it can acquire a net negative charge when the bee is near. The stigma’s attraction to the pollen is therefore of the
monopole–dipole type, whereas the anther’s attraction to the pollen is of the smaller dipole–dipole type.

The ability of the gecko lizard to stick to a window or walk upside down on a ceiling is due to the van der Waals force. This force is an interaction between dipoles in the gecko’s feet and dipoles on the surface. The force falls off rapidly with distance, being proportional to $1/r^7$. Equivalently, the potential energy is proportional to $1/r^6$. (In short, quantum fluctuations create random dipole moments in a given molecule. The resulting $E \propto 1/r^3$ field induces a dipole moment $p \propto E \propto 1/r^3$ in a neighboring molecule. The resulting $(\mathbf{p} \cdot \nabla)E$ force on this molecule is then proportional to $1/r^7$.) The key for geckos, therefore, is to make the distance $r$ be as small as possible. They do this by having hundreds of thousands of tiny hairs (called setae) on their feet, each of which contains hundreds of even tinier hairs called spatulae. These spatulae are able to penetrate the nooks and crannies on the surface, making the various distances $r$ extremely small. If all of a gecko’s spatulae are engaged, it could walk on a ceiling with a few hundred pounds strapped to it!

The main ingredient in soaps and detergents is a surfactant. This is a molecule in the form of a long chain with a polar end and a nonpolar end. The polar end is attracted to the polar water molecules (it is called hydrophilic), whereas the nonpolar end isn’t (it is called hydrophobic). The nonpolar end is instead attracted to other hydrophobic molecules, such as oils and other grime on you or your clothes. More precisely, the hydrophobic ends/molecules aren’t actually attracted to each other. Rather, the attraction of all the polar molecules to themselves has the effect of forcing all the hydrophobic ends/molecules into little clumps, called micelles. This gives the appearance of an attraction. (The same reasoning leads to the everyday fact that oil and water don’t mix.) The clumps (with the oils and nonpolar surfactant ends in the interior, and the polar surfactant ends on the surface) float around in the water and can be eliminated by discarding the water, that is, by rinsing. So your laundry detergent won’t work without water!

The large dipole moment of a water molecule is what allows a microwave oven to heat your food. The alternating electric field of the microwave radiation (created by a magnetron; see Section 8.7) causes the water dipoles to rotate back and forth. This jiggling of the molecules causes them to bump into each other, which results in the thermal energy (the heat) that you observe. The specific microwave frequency that is used (usually about 2.5 GHz, which corresponds to a wavelength of about 12 cm) has nothing to do with the resonant vibrational frequency of a free water molecule in vapor, which is about 20 GHz. There isn’t anything special about the 12 cm wavelength, although if it were much shorter the waves wouldn’t penetrate the food as well. The water molecules in ice can’t rotate as easily, so that’s why it takes a while to defrost frozen food. You can’t just crank up the power because, if one part of the food thaws first, then it will absorb energy much faster than
the remaining frozen part. You will then end up with, for example, thor-
oughly cooked meat right next to frozen meat. The defrost cycle in a
microwave oven functions by simply shutting off for periods of time,
allowing the heat to diffuse.

When some materials, such as quartz, are put under stress and bent,
a voltage difference arises between different parts. This is known as the
piezoelectric effect. (Conversely, if a voltage difference is applied to dif-
ferent parts, the material will bend.) What happens is that the molecules
are stretched or squashed, and for certain configurations this results in the
molecules acquiring a dipole moment. The material therefore behaves
like a polarized dielectric; there are net surface charges, so the result is
effectively a capacitor with a voltage difference between the plates. The
piezoelectric property of quartz allows your wristwatch to keep time.
A tiny quartz crystal, which is the analog of the pendulum in a pendulum
clock, is cut so that it vibrates with a specific resonant frequency, usually
$2^{15} = 32,768$ Hz. (This power of 2 makes it easy for a chain of frequency
dividers to generate the desired 1 Hz frequency.) Via the piezoelectric
effect, an electric signal with this 32,768 Hz frequency is sent to a cir-
cuit. The circuit then amplifies the signal and sends it back to the crystal,
providing the necessary driving force to keep the resonant oscillation
going. Quartz has a very high $Q$ value, so only a tiny amount of input
energy is needed. That is why the battery can last so long. The oscilla-
tions are initially produced by the random ac noise in the circuit. This
ac noise contains at least a little bit of the resonant frequency, which the
crystal responds to.

**CHAPTER SUMMARY**

- If a capacitor is filled with a dielectric (an insulator), the capacitance
  increases by a factor $\kappa$, known as the dielectric constant. This is a
  consequence of the fact that the polarization of the molecules in the
dielectric causes layers of charge to form near the capacitor plates,
partially canceling the free charge.

- The potential due to a charge distribution can be written as the sum of
terms with increasing powers of $1/r$. The coefficients of these terms
  are called moments. A net charge has a monopole moment. Two oppo-
site monopoles create a dipole. Two opposite dipoles create a quadru-
pole, and so on. The potential and field due to a dipole are given by

$$
\phi(r, \theta) = \frac{p \cos \theta}{4 \pi \epsilon_0 r^2},
E(r, \theta) = \frac{p}{4 \pi \epsilon_0 r^3} \left(2 \cos \theta \mathbf{\hat{r}} + \sin \theta \mathbf{\hat{\theta}}\right).
$$

(10.84)

- The torque on an electric dipole is $\mathbf{N} = \mathbf{p} \times \mathbf{E}$. The force is $F_x = \mathbf{p} \cdot \nabla E_x$, and likewise for the $y$ and $z$ components.
• An external electric field will cause an atom to become polarized. The *
atomic polarizability* $\alpha$ is defined by $p = \alpha E$. However, the quantity $\alpha/4\pi\varepsilon_0$, with has the dimensions of volume, is also commonly called the atomic polarizability. In order of magnitude, $\alpha/4\pi\varepsilon_0$ equals an atomic volume.

• Some molecules have a permanent dipole moment; the moment exists even in the absence of an external electric field. These are called *polar* molecules. An external electric field causes the dipoles to align (at least partially), which leads to an overall polarization of the material.

• The *polarization* per unit volume is given by $P = pN$. Uniformly polarized matter is equivalent to a surface charge density $\sigma = P$, or $\sigma = P\cos\theta$ if the surface is tilted with respect to the direction of $P$.

• When we talk about the electric field inside matter, we mean the spatial average, $\langle E \rangle = \left(1/V\right)\int E\,dv$. Inside a uniformly polarized slab, this average is $-P/\varepsilon_0$.

• The *electric susceptibility* $\chi_e$ is defined by

$$\chi_e \equiv \frac{P}{\varepsilon_0 E} \implies \chi_e = \kappa - 1. \quad (10.85)$$

• The field inside a uniformly polarized sphere is $-P/3\varepsilon_0$. If a dielectric sphere is placed in a uniform electric field $E_0$, the resulting polarization is uniform and is given by $P = 3\varepsilon_0(\kappa - 1)E_0/(\kappa + 2)$.

• For any material, $\text{div } P = -\rho_{\text{bound}}$. Combining this with Gauss’s law, $\text{div } E = \rho_{\text{total}}/\varepsilon_0$, gives

$$\text{div } D = \rho_{\text{free}}, \quad \text{where } D \equiv \varepsilon_0 E + P. \quad (10.86)$$

$D$ is known as the *electric displacement vector*. If additionally we are dealing with a linear dielectric, then

$$D \equiv \varepsilon E, \quad \text{where } \varepsilon \equiv \kappa\varepsilon_0. \quad (10.87)$$

• If an external electric field is applied to a dielectric containing polar molecules, the polarizations tend to align with the field, but thermal energy generally prevents the alignment from being large. The susceptibility is given roughly by $\chi_e \approx Np^2/\varepsilon_0 kT$.

• In a rapidly changing electric field, the induced polarization of atoms and molecules can keep up with the field at high frequencies. However, the polarization arising from polar molecules cannot, because it is much more difficult to rotate a molecule as a whole than simply to stretch it.

• The *bound-charge* current density satisfies $J_{\text{bound}} = dP/dt$. This represents a true current, but when writing the “curl $B$” Maxwell equation it is often convenient to separate this current from the free-charge current:

$$\text{curl } B = \mu_0 \left(\frac{\varepsilon_0}{\varepsilon} \frac{\partial E}{\partial t} + \frac{\partial P}{\partial t} + J_{\text{free}}\right) \equiv \mu_0 \left(\frac{\partial D}{\partial t} + J_{\text{free}}\right). \quad (10.88)$$
Problems

10.1 Leaky cell membrane  
In Section 4.11 we discussed the relaxation time of a capacitor filled with a material having a resistivity $\rho$. If you look back at that discussion you will notice that we dodged the question of the dielectric constant of the material. Now you can repair that omission, by introducing $\kappa$ properly into the expression for the time constant. A leaky capacitor important to us all is formed by the wall of a living cell, an insulator (among its many other functions!) that separates two conducting fluids. Its electrical properties are of particular interest in the case of the nerve cell, for the propagation of a nerve impulse is accompanied by rapid changes in the electric potential difference between interior and exterior.

(a) The cell membrane typically has a capacitance around 1 microfarad per square centimeter of membrane area. It is believed the membrane consists of material having a dielectric constant about 3. What thickness does this imply?

(b) Other electrical measurements have indicated that the resistance of 1 cm$^2$ of cell membrane, measured from the conducting fluid on one side to that on the other, is around 1000 ohms. Show that the time constant of such a leaky capacitor is independent of the area of the capacitor. How large is it in this case? Where would the resistivity $\rho$ of such membrane material fall on the chart of Fig. 4.8?

10.2 Force on a dielectric  
A rectangular capacitor with side lengths $a$ and $b$ has separation $s$, with $s$ much smaller than $a$ and $b$. It is partially filled with a dielectric with dielectric constant $\kappa$. The overlap distance is $x$; see Fig. 10.31. The capacitor is isolated and has constant charge $Q$.

(a) What is the energy stored in the system? (Treat the capacitor like two capacitors in parallel.)

(b) What is the force on the dielectric? Does this force pull the dielectric into the capacitor or push it out?

10.3 Energy of dipoles  
Find the potential energy of the first and third dipole configurations in Exercise 10.29 (the second and fourth require only a slight modification), by explicitly looking at the potential energy of the
Electric fields in matter

10.4 Dipole polar components **
Show that Eq. (10.18) follows from Eq. (10.17). Hint: You can write the Cartesian unit vectors in terms of the polar unit vectors, or you can project the vector \((E_x, E_z)\) onto the radial and tangential directions.

10.5 Average field **
(a) (This problem builds on the results from Problem 1.28.) Given an arbitrary collection of charges inside a sphere of radius \(R\), show that the average electric field over the volume of the sphere is given by \(E_{\text{avg}} = -p/4\pi\epsilon_0 R^3\), where \(p\) is the total dipole moment, measured relative to the center.
(b) For the specific case of the dipole shown in Fig. 10.32(a), find the average electric field over both the surface and the volume of the sphere of radius \(R\).
(c) Repeat for the case shown in Fig. 10.32(b).

10.6 Quadrupole tensor ***
You should see the quadrupole tensor at least once in your life, so here it is. Calculate the general form of the quadrupole tensor by writing the \(1/r^3\) part of the potential in a form that has all of its dependence on the primed coordinates collected into a matrix. (By analogy, the \(1/r^2\) part of the potential in Eq. (10.12) has all of its dependence on the primed coordinates collected into the vector \(p\) given by Eq. (10.13).)

10.7 Force on a dipole **
Derive Eq. (10.26). As usual, work in the approximation where the dipole length \(s\) is small. Hint: Let the two charges be at positions \(r\) and \(r + s\).

10.8 Force from an induced dipole **
Between any ion and any neutral atom there is a force that arises as follows. The electric field of the ion polarizes the atom; the field of that induced dipole reacts on the ion. Show that this force is always attractive, and that it varies with the inverse fifth power of the distance of separation \(r\). Derive an expression for the associated
potential energy, with zero energy corresponding to infinite separation. For what distance \( r \) does this potential energy have the same magnitude as \( kT \) at room temperature (which is \( 4 \cdot 10^{-21} \) joule) if the ion is singly charged and the atom is a sodium atom? (See Table 10.2.)

10.9 Polarized water  
The electric dipole moment of the water molecule is given in Fig. 10.14. Imagine that all the molecular dipoles in a cup of water could be made to point down. Calculate the magnitude of the resulting surface charge density at the upper water surface, and express it in electrons per square centimeter.

10.10 Tangent field lines  
Assume that the uniform field \( \mathbf{E}_0 \) that causes the electric field in Fig. 10.27 is produced by large capacitor plates very far away. Consider the special set of field lines that are tangent to the sphere. These lines hit each of the distant capacitor plates in a circle of radius \( r \). What is \( r \) in terms of the radius \( R \) and dielectric constant \( \kappa \) of the sphere? Hint: Consider a well-chosen Gaussian surface that has the horizontal great circle of the sphere as part of its boundary.

10.11 Bound charge and divergence of \( \mathbf{P} \)  
Derive Eq. (10.61) by considering the volume integral of both sides of the equation. Assume that the dipoles consist of charges \( \pm q \) separated by a distance \( s \). Hint: Consider a small patch of the surface bounding a given volume. What causes there to be a net bound charge inside the volume due to the dipoles near this patch?

10.12 Boundary conditions on \( \mathbf{D} \)  
Using \( \mathbf{D} \equiv \varepsilon_0 \mathbf{E} + \mathbf{P} \) and \( \text{div} \, \mathbf{D} = \rho_{\text{free}} \), derive the general rules for the discontinuities (if any) in \( \mathbf{D}_\parallel \) and \( \mathbf{D}_\perp \) across the surface of an arbitrarily shaped polarized material (with no free charge).

10.13 \( Q \) for a leaky capacitor  
Consider an oscillating electric field, \( E_0 \cos \omega t \), inside a dielectric medium that is not a perfect insulator. The medium has dielectric constant \( \kappa \) and conductivity \( \sigma \). This could be the electric field in some leaky capacitor that is part of a resonant circuit, or it could be the electric field at a particular location in an electromagnetic wave. Show that the \( Q \) factor, as defined by Eq. (8.12), is \( \omega \varepsilon / \sigma \) for this system, and evaluate it for seawater at a frequency of 1000 MHz. You will need to use the result from Exercise 10.42. (The conductivity is given in Table 4.1, and the dielectric constant may be assumed to be the same as that of pure water at the same frequency. See Fig. 10.29.) What does your result suggest about the propagation of decimeter waves through seawater?
10.14Boundary conditions on $E$ and $B$ **
Find the boundary conditions on $E_\parallel$, $E_\perp$, $B_\parallel$, and $B_\perp$ across the interface between two linear dielectrics. Assume that there are no free charges or free currents present.

Exercises

10.15Densities on a capacitor **
Consider the setup of Problem 10.2. In terms of the various parameters given there, find the charge densities on the left and right parts of the capacitor. You should find that as $x$ increases, the charge densities on both parts of plates decrease. At first glance this seems a bit absurd, so try to explain intuitively how it is possible.

10.16Leyden jar **
In 1746 a Professor Musschenbroek in Leiden charged water in a bottle by touching a wire, projecting from the neck of the bottle, to his electrostatic machine. When his assistant, who was holding the bottle in one hand, tried to remove the wire with the other, he got a violent shock. Thus did the simple capacitor force itself on the attention of electrical scientists. The discovery of the “Leyden jar” revolutionized electrical experimentation. In 1747 Benjamin Franklin was already writing about his experiments with “Mr. Musschenbroek’s wonderful bottle.” The jar was really nothing but glass with a conductor on each side of it. To see why it caused such a sensation, estimate roughly the capacitance of a jar made of a 1 liter bottle with walls 2 mm thick, the glass having a dielectric constant 4. What diameter sphere, in air, would have the same capacitance?

10.17Maximum energy storage **
Materials to be used as insulators or dielectrics in capacitors are rated with respect to dielectric strength, defined as the maximum internal electric field the material can support without danger of electrical breakdown. It is customary to express the dielectric strength in kilovolts per mil. (One mil is 0.001 inch, or 0.00254 cm.) For example, Mylar (a Dupont polyester film) is rated as having a dielectric strength of 14 kilovolts/mil when it is used in a thin sheet – as it would be in a typical capacitor. The dielectric constant $\kappa$ of Mylar is 3.25. Its density is 1.40 g/cm$^3$. Calculate the maximum amount of energy that can be stored in a Mylar-filled capacitor, and express it in joules/kg of Mylar. Assuming the electrodes and case account for 25 percent of the capacitor’s weight, how high could the capacitor be lifted by the energy stored in it? Compare the capacitor as an energy storage device with the battery in Exercise 4.41.
10.18  **Partially filled capacitors**

Figure 10.33 shows three capacitors of the same area and plate separation. Call the capacitance of the vacuum capacitor \(C_0\). Each of the others is half-filled with a dielectric, with the same dielectric constant \(\kappa\), but differently disposed, as shown. Find the capacitance of each of these two capacitors. (Neglect edge effects.)

10.19  **Capacitor roll**

You have a supply of polyethylene tape, with dielectric constant 2.3, that is 2.25 inches wide and 0.001 inch thick; you also have a supply of aluminum tape that is 2 inches wide and 0.0005 inch thick. You want to make a capacitor of about 0.05 microfarad capacitance, in the form of a compact cylindrical roll. Describe how you might do this, estimating the amount of tape of each kind that would be needed, and the overall diameter of the finished capacitor. (It may help to look at Problem 3.21 and Exercise 3.57.)

10.20  **Work in a dipole field**

How much work is done in moving unit positive charge from \(A\) to \(B\) in the field of the dipole \(p\) shown in Fig. 10.34?

10.21  **A few dipole moments**

What is the magnitude and direction of the dipole moment vector \(p\) of each of the charge distributions in parts (a), (b), and (c) of Fig. 10.35?

10.22  **Fringing field from a capacitor**

A parallel-plate capacitor, with a measured capacitance \(C = 250\) picofarads \((250 \cdot 10^{-12} \text{ F})\), is charged to a potential difference of 2000 volts. The plates are 1.5 cm apart. We are interested in the field outside the capacitor, the “fringing” field which we usually ignore. In particular, we would like to know the field at a distance from the capacitor large compared with the size of the capacitor itself. This can be found by treating the charge distribution on the capacitor as a dipole. Estimate the electric field strength

(a) at a point 3 meters from the capacitor in the plane of the plates;
(b) at a point the same distance away, in a direction perpendicular to the plates.

10.23  **Dipole field plus uniform field**

A dipole of strength \(p = 6 \cdot 10^{-10} \text{ C-m}\) is located at the origin, pointing in the \(\hat{z}\) direction. To its field is added a uniform electric field of strength 150 kV/m in the \(\hat{y}\) direction. At how many places, located where, is the total field zero?

10.24  **Field lines**

A field line in the dipole field is described in polar coordinates by the very simple equation \(r = r_0 \sin^2 \theta\), in which \(r_0\) is the radius
at which the field line passes through the equatorial plane of the dipole. Show that this is true by demonstrating that at any point on that curve the tangent has the same direction as the dipole field.

10.25 *Average dipole field on a sphere* **
By direct integration, show that the average of the dipole field, over the surface of a sphere centered at the dipole, is zero. You will want to work with the Cartesian components of \( E \), but feel free to write these components in terms of spherical coordinates, as in Eq. (10.17).

10.26 *Quadrupole for a square* **
Calculate the quadrupole matrix \( Q \) (see the result of Problem 10.6) for the configuration of charges in Fig. 10.5. Then show that the potential at a point at (large) radius \( r \) on the \( z \equiv x_3 \) axis equals \( 3ea^2/4\pi\epsilon_0r^3 \), where \( a \) is the distance from all the charges to the origin. What is the potential at the point \( (r/\sqrt{2}, 1, 0, 1) \)?

10.27 *Pascal’s triangle and the multipole expansion* **
(a) If two monopoles with opposite sign are placed near each other, they make a dipole. Likewise, if two dipoles with opposite sign are placed near each other, they make a quadrupole, and so on. Explain how this fact was used in Fig. 10.36 to obtain each of the configurations from the one above it. Note that the magnitudes of the charges form Pascal’s triangle. This triangle provides a very simple way of generating successively higher terms in the multipole expansion.

(b) If the bottom configuration in Fig. 10.36 is indeed an octupole, then the leading-order term in the potential at the point \( P \) in Fig. 10.37 must be of order \( 1/r^4 \). (This is two orders of \( 1/r \) higher than the \( 1/r^2 \) dipole potential.) Verify this. That is, use a Taylor series to show that the order 1, \( 1/r \), \( 1/r^2 \), and \( 1/r^3 \) terms in the potential vanish. Define \( r \) to be the distance to the rightmost charge, and assume \( r \gg a \). This problem is easy if you use the \texttt{Series} operation in \textit{Mathematica}, but you should work it out by hand. It is interesting to see how each of the terms vanishes.

As you will discover, this result is related to a nice little theorem about the sum \( \sum_{k=0}^{N} \binom{N}{k}m^k(-1)^k \), where \( m \) can take on any value from 0 to \( N - 1 \). You are encouraged to think about this, although you don’t need to prove it here. \textit{Hint:} Expand \( (1 - x)^N \) with the binomial theorem, and take the derivative. Then multiply by \( x \) and take the derivative again. Repeat this process as needed, and then set \( x = 1 \).

10.28 *Force on a dipole* **
What are the magnitude and direction of the force on the central dipole caused by the field of the other two dipoles in Fig. 10.38?
10.29 Energy of dipole pairs **

Shown in Fig. 10.39 are four different arrangements of the electric dipole moments of two neighboring polar molecules. Find the potential energy of each arrangement, the potential energy being defined as the work done in bringing the two molecules together from infinite separation while keeping their moments in the specified orientation. That is not necessarily the easiest way to calculate it. You can always bring them together one way and then rotate them.

10.30 Polarized hydrogen **

A hydrogen atom is placed in an electric field $E$. The proton and the electron cloud are pulled in opposite directions. Assume simplistically (since we are concerned only with a rough result here) that the electron cloud takes the form of a uniform sphere with radius $a$, with the proton a distance $\Delta z$ from the center, as shown in Fig. 10.40. Find $\Delta z$, and show that your result agrees with Eq. (10.27).

10.31 Mutually induced dipoles **

Two polarizable atoms $A$ and $B$ are a fixed distance apart. The polarizability of each atom is $\alpha$. Consider the following intriguing possibility. Atom $A$ is polarized by an electric field, the source of which is the electric dipole moment $p_B$ of atom $B$. This dipole moment is induced in atom $B$ by an electric field, the source of which is the dipole moment $p_A$ of atom $A$. Can this happen? If so, under what conditions? If not, why not?

10.32 Hydration *

The phenomenon of hydration is important in the chemistry of aqueous solutions. This refers to the fact that an ion in solution gathers around itself a cluster of water molecules, which cling to it rather tightly. The force of attraction between a dipole and a point charge is responsible for this. Estimate the energy required to separate an ion carrying a single charge $e$ from a water molecule, assuming that initially the ion is located 1.5 angstroms from the effective location of the $\text{H}_2\text{O}$ dipole. (This distance is actually a rather ill-defined quantity, since the water molecule, viewed from close up, is a charge distribution, not an infinitesimal dipole.) Which part of the water molecule will be found nearest to a negative ion? See Fig. 10.14 for the dipole moment of the water molecule.
10.33 *Field from hydrogen chloride*

A hydrogen chloride molecule is located at the origin, with the H–Cl line along the z axis and Cl uppermost. What is the direction of the electric field, and its strength in volts/m, at a point 10 angstroms up from the origin, on the z axis? At a point 10 angstroms out from the origin, on the y axis? (p is given in Fig. 10.14.)

10.34 **Hydrogen chloride dipole moment**

In the hydrogen chloride molecule the distance between the chlorine nucleus and the proton is 1.28 angstroms. Suppose the electron from the hydrogen atom is transferred entirely to the chlorine atom, joining with the other electrons to form a spherically symmetrical negative charge that is centered on the chlorine nucleus. How does the electric dipole moment of this model compare with the actual HCl dipole moment given in Fig. 10.14? Where must the actual “center of gravity” of the negative charge distribution be located in the real molecule? (The chlorine nucleus has a charge 17e, and the hydrogen nucleus has a charge e.)

10.35 **Some electric susceptibilities**

From the values of κ given for water, ammonia, and methanol in Table 10.1, we know that the electric susceptibility \( \chi_e \) for each liquid is given by \( \chi_e = \kappa - 1 \). Our theoretical prediction in Eq. (10.73) can be written \( \chi_e = C n p^2 / \epsilon_0 kT \), with the factor C as yet unknown, but expected to have order of magnitude unity. The densities of the liquids are 1.00, 0.82, and 1.33 g/cm\(^3\), respectively; their molecular weights are 18, 17, and 32. Taking the value of the dipole moment from Fig. 10.14, find for each case the value of C required to fit the observed value of \( \chi_e \).

10.36 **Discontinuity in \( E_\perp \)**

Consider the polarized sphere from Section 10.9. Using the forms of the internal and external electric fields, show that \( E_\perp \) has a discontinuity of \( P_\perp / \epsilon_0 = P \cos \theta / \epsilon_0 \) everywhere on the surface of the sphere.

10.37 **E at the center of a polarized sphere**

If you don’t trust the \( E = -P/3\epsilon_0 \) result we obtained in Section 10.9 for the field inside a uniformly polarized sphere, you will find it more believable if you check it in a special case. By direct integration of the contributions from the \( \sigma = P \cos \theta \) surface charge density, show that the field at the center is directed downward (assuming \( \mathbf{P} \) points upward) with magnitude \( P/3\epsilon_0 \).

10.38 **Uniform field via superposition**

In Section 10.9, the fact that the electric field is uniform inside the polarized sphere was deduced from the form of the potential on the
boundary. You can also prove it by superposing the internal fields of two balls of charge whose centers are separated.

(a) Show that, inside a spherical uniform charge distribution, \( E \) is proportional to \( r \).

(b) Now take two spherical distributions with density \( \rho \) and \( -\rho \), centers at \( C_1 \) and \( C_2 \), and show that the resultant field is constant and parallel to the line from \( C_1 \) to \( C_2 \). Verify that the field can be written as \( -P/3\varepsilon_0 \).

(c) Analyze in the same way the field of a long cylindrical rod that is polarized perpendicular to its axis.

10.39 **Conducting-sphere limit**

Our formula for the dielectric sphere in Section 10.10 can actually serve to describe a metal sphere in a uniform field. To demonstrate this, investigate the limiting case \( \kappa \to \infty \), and show that the external field then takes on a form that satisfies the perfect-conductor boundary conditions. What about the internal field? Make a sketch of some field lines for this limiting case. What is the radius of a conducting sphere with polarizability equal to that of the hydrogen atom, given in Table 10.2?

10.40 **Continuity of \( D \)**

Use the definition of \( D \), namely \( D \equiv \varepsilon_0 E + P \), to show that \( D \) is continuous across the faces of a uniformly polarized slab. Assume that the polarization is perpendicular to the faces, and that the thickness of the slab is small compared with the other two dimensions.

10.41 **Discontinuity in \( D_\parallel \)**

Consider the polarized sphere from Section 10.9. Using the forms of the internal and external electric fields, find the discontinuity in \( D_\parallel \) across the surface of the sphere, as a function of \( \theta \).

10.42 **Energy density in a dielectric**

By considering how the introduction of a dielectric changes the energy stored in a capacitor, show that the correct expression for the energy density in a dielectric must be \( \varepsilon E^2/2 \). Then compare the energy stored in the electric field with that stored in the magnetic field in the wave studied in Section 10.15.

10.43 **Reflected wave**

A block of glass, refractive index \( n = \sqrt{\kappa} \), fills the space \( y > 0 \), its surface being the \( xz \) plane. A plane wave traveling in the positive \( y \) direction through the empty space \( y < 0 \) is incident upon this surface. The electric field in this wave is \( \hat{z}E_i \sin(ky - \omega t) \). There is a transmitted wave inside the glass block, with an electric field given by \( \hat{z}E_i \sin(k'y - \omega t) \). There is also a reflected wave in the space \( y < 0 \), traveling away from the glass in the negative \( y \) direction.
Its electric field is $\hat{E}_r \sin(ky + \omega t)$. Of course, each wave has its magnetic field, of amplitude, respectively, $B_i$, $B_t$, and $B_r$.

The total magnetic field must be continuous at $y = 0$; and the total electric field, being parallel to the surface, must be continuous also (see Problem 10.14). Show that these requirements, and the relation of $B_t$ to $E_t$ given in Eq. (10.83) (with the “0” subscript changed to “t”), suffice to determine the ratio of $E_r$ to $E_i$. When a light wave is incident normally at an air–glass interface, what fraction of the energy is reflected if the index $n$ is 1.6?
Magnetic fields in matter are a bit more involved than electric fields in matter. Our main goal in this chapter is to understand the three types of magnetic materials: diamagnetic materials, which are weakly repelled by a solenoid; paramagnetic materials, which are somewhat strongly attracted; and ferromagnetic materials, which are very strongly attracted. As was the case in Chapter 10, we will need to understand dipoles. The far field of a magnetic dipole has the same form as that of an electric dipole, with the magnetic dipole moment replacing the electric dipole moment. However, the near fields are fundamentally different due to the absence of magnetic charge. We will find that diamagnetism is due to the fact that an applied magnetic field causes the magnetic dipole moment arising from the orbital motion of electrons in atoms to pick up a contribution pointing opposite to the applied field. In contrast, in the case of paramagnetism, the spin dipole moment is the relevant one, and it picks up a contribution pointing in the same direction as the applied field. Ferromagnetism is similar to paramagnetism, although a certain quantum phenomenon makes the overall effect much larger; a ferromagnetic dipole moment can exist in the absence of an external magnetic field. Magnetized materials can be described by the magnetization $M$, the curl of which gives the bound currents (which arise from both orbital motion and spin). By considering separately the free and bound currents, we are led to the field $H$ (also called the “magnetic field”) whose curl involves only the free current (unlike the magnetic field $B$, whose curl involves all the current, by Ampère’s law).
11.1 How various substances respond to a magnetic field

Imagine doing some experiments with a very intense magnetic field. To be definite, suppose we have built a solenoid of 10 cm inside diameter, 40 cm long, like the one shown in Fig. 11.1. Its outer diameter is 40 cm, most of the space being filled with copper windings. This coil will provide a steady field of 3.0 tesla, or 30,000 gauss, at its center if supplied with 400 kilowatts of electric power – and something like 30 gallons of water per minute, to carry off the heat. We mention these practical details to show that our device, though nothing extraordinary, is a pretty respectable laboratory magnet. The field strength at the center is nearly $10^5$ times the earth’s field, and probably 5 or 10 times stronger than the field near any iron bar magnet or horseshoe magnet you may have experimented with, although some rare-earth magnets can have fields of around 1 tesla.

The field will be fairly uniform near the center of the solenoid, falling, on the axis at either end, to roughly half its central value. It will be rather less uniform than the field of the solenoid in Fig. 6.18, since our coil is equivalent to a “nested” superposition of solenoids with length–diameter ratio varying from 4:1 to 1:1. In fact, if we analyze our coil in that way and use the formula that we derived for the field on the axis of a solenoid with a single-layer winding (see Eq. (6.56)), it is not hard to calculate the axial field exactly. A graph of the field strength on the axis, with the central field taken as 3.0 tesla = 30 kilogauss, is included in Fig. 11.1. The intensity just at the end of the coil is 1.8 tesla, and in
that neighborhood the field is changing with a gradient of approximately 17 tesla/m, or 1700 gauss/cm.

Let’s put various substances into this field and see if a force acts on them. Generally, we do detect a force. It vanishes when the current in the coil is switched off. We soon discover that the force is strongest not when our sample of substance is at the center of the coil where the magnetic field $B_z$ is strongest, but when it is located near the end of the coil where the gradient $dB_z/dz$ is large. From now on let us support each sample just inside the upper end of the coil. Figure 11.2 shows one such sample, contained in a test tube suspended by a spring which can be calibrated to indicate the extra force caused by the magnetic field. Naturally we have to do a “blank” experiment with the test tube and suspension alone, to allow for the magnetic force on everything other than the sample.

We find in such an experiment that the force on a particular substance – metallic aluminum, for instance – is proportional to the mass of the sample and independent of its shape, as long as the sample is not too large. (Experiments with a small sample in this coil show that the force remains practically constant over a region a few centimeters in extent, inside the end of the coil; if we use samples no more than 1 to 2 cm$^3$ in volume, they can be kept well within this region.) We can express our quantitative results, for a given substance, as so many newtons force per kilogram of sample, under the conditions $B_z = 1.8$ tesla, $dB_z/dz = 17$ tesla/m.

But first the qualitative results, which are a bit bewildering. For a large number of quite ordinary pure substances, the force observed, although easily measurable, seems ridiculously small, despite all our effort to provide an intense magnetic field. Typically, the force is 0.1 or 0.2 newtons per kilogram, that is, no more than a few percent of the weight of the sample (which is 9.8 newtons per kilogram). It is directed upward for some samples, downward for others. This has nothing to do with the direction of the magnetic field, as we can verify by reversing the current in the coil. Instead, it appears that some substances are always pulled in the direction of increasing field intensity, others in the direction of decreasing field intensity, irrespective of the field direction.

We do find some substances that are attracted to the coil with considerably greater force. For instance, copper chloride crystals are pulled downward with a force of 2.8 newtons per kilogram of sample. Liquid oxygen behaves spectacularly in this experiment; it is pulled into the coil with a force nearly eight times its weight. In fact, if we were to bring an uncovered flask of liquid oxygen up to the bottom end of our coil, the liquid would be lifted right out of the flask. (Where do you think it would end up?) Liquid nitrogen, on the other hand, proves to be quite unexciting; it is pushed away from the coil with the feeble force of 0.1 newtons per kilogram. In Table 11.1 we have listed some results that one might obtain in such an experiment. The substances, including those
Magnetic fields in matter

Table 11.1.
Force per kilogram near the upper end of the coil in our experiment, where $B_z = 1.8$ tesla and $dB_z/dz = 17$ tesla/m

<table>
<thead>
<tr>
<th>Substance</th>
<th>Formula</th>
<th>Force (newtons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diamagnetic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>H$_2$O</td>
<td>$-0.22$</td>
</tr>
<tr>
<td>Copper</td>
<td>Cu</td>
<td>$-0.026$</td>
</tr>
<tr>
<td>Sodium chloride</td>
<td>NaCl</td>
<td>$-0.15$</td>
</tr>
<tr>
<td>Sulfur</td>
<td>S</td>
<td>$-0.16$</td>
</tr>
<tr>
<td>Diamond</td>
<td>C</td>
<td>$-0.16$</td>
</tr>
<tr>
<td>Graphite</td>
<td>C</td>
<td>$-1.10$</td>
</tr>
<tr>
<td>Liquid nitrogen</td>
<td>N$_2$</td>
<td>$-0.10$ (78 K)</td>
</tr>
<tr>
<td>Paramagnetic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium</td>
<td>Na</td>
<td>$0.20$</td>
</tr>
<tr>
<td>Aluminum</td>
<td>Al</td>
<td>$0.17$</td>
</tr>
<tr>
<td>Copper chloride</td>
<td>CuCl$_2$</td>
<td>$2.8$</td>
</tr>
<tr>
<td>Nickel sulfate</td>
<td>NiSO$_4$</td>
<td>$8.3$</td>
</tr>
<tr>
<td>Liquid oxygen</td>
<td>O$_2$</td>
<td>$75$ (90 K)</td>
</tr>
<tr>
<td>Ferromagnetic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron</td>
<td>Fe</td>
<td>$4000$</td>
</tr>
<tr>
<td>Magnetite</td>
<td>Fe$_3$O$_4$</td>
<td>$1200$</td>
</tr>
</tbody>
</table>

Direction of force: downward (into coil), +; upward, −.
All measurements were made at a temperature of 20°C unless otherwise stated.
The three types of magnetism are defined in the text.

already mentioned, have been chosen to suggest, as best one can with a sparse sampling, the wide range of magnetic behavior we find in ordinary materials. Note that our convention for the sign of the force is that a positive force is directed into the coil.

As you know, a few substances, of which the most familiar is metallic iron, seem far more “magnetic” than any others. In Table 11.1 we give the force per kilogram that would act on a piece of iron put in the same position in the field as the other samples. Since 1 newton is about 0.22 pounds, the force per kilogram is roughly 900 pounds, or nearly 1 pound for a 1 gram sample! (We would not have been so naive as to approach our magnet with a gram of iron suspended in a test tube from a delicate spring – a different suspension would have to be used.) Observe that there is a factor of more than $10^5$ between the force per kilogram on iron and the force per kilogram on copper, elements not otherwise radically different. Incidentally, this suggests that reliable magnetic measurements on a substance like copper may not be easy. A few parts per million contamination by metallic iron particles would utterly falsify the result.

There is another essential difference between the behavior of the iron and the magnetite and that of the other substances in the table. Suppose we make the obvious test, by varying the field strength of the magnet, to ascertain whether the force on a sample is proportional to the
11.1 Substances responding to magnetic field

field. For instance, we might reduce the solenoid current by half, thereby halving both the field intensity $B_z$ and its gradient $dB_z/dz$. We would find, in the case of every substance above iron in the table, that the force is reduced to one-fourth its former value, whereas the force on the iron sample, and that on the magnetite, would be reduced only to one-half or perhaps a little less. Evidently the force, under these conditions at least, is proportional to the square of the field strength for all the other substances listed, but nearly proportional to the field strength itself for Fe and Fe$_3$O$_4$.

It appears that we may be dealing with several different phenomena here, and complicated ones at that. As a small step toward understanding, we can introduce some classification.

(1) Diamagnetism First, those substances that are feebly repelled by our magnet – water, sodium chloride, diamond, etc. – are called diamagnetic. The majority of inorganic compounds and practically all organic compounds are diamagnetic. It turns out, in fact, that diamagnetism is a property of every atom and molecule. When the opposite behavior is observed, it is because the diamagnetism is outweighed by a different and stronger effect, one that leads to attraction.

(2) Paramagnetism Substances that are attracted toward the region of stronger field are called paramagnetic. In some cases, notably metals such as aluminum, sodium, and many others, the paramagnetism is not much stronger than the common diamagnetism. In other materials, such as the NiSO$_4$ and the CuCl$_2$ on our list, the paramagnetic effect is much stronger. In these substances also, it increases as the temperature is lowered, leading to quite large effects at temperatures near absolute zero. The increase of paramagnetism with lowering temperature is responsible in part for the large force recorded for liquid oxygen. If you think all this is going to be easy to explain, observe that copper is diamagnetic while copper chloride is paramagnetic, but sodium is paramagnetic while sodium chloride is diamagnetic.

(3) Ferromagnetism Finally, substances that behave like iron and magnetite are called ferromagnetic. In addition to the common metals of this class – iron, cobalt, and nickel – quite a number of ferromagnetic alloys and crystalline compounds are known. Indeed current research in ferromagnetism is steadily lengthening the list.

In this chapter we have two tasks. One is to develop a treatment of the large-scale phenomena involving magnetized matter, in which the material itself is characterized by a few parameters and the experimentally determined relations among them. It is like a treatment of dielectrics based on some observed relation between electric field and bulk polarization. We sometimes call such a theory phenomenological; it is more of a description than an explanation. Our second task is to try to understand, at least in a general way, the atomic origin of the various magnetic
effects. Even more than dielectric phenomena, the magnetic effects, once understood, reveal some basic features of atomic structure.

One general fact stands out in Table 11.1. Very little energy, on the scale of molecular energies, is involved in diamagnetism and paramagnetism. Take the extreme example of liquid oxygen. To pull 1 kilogram (although the sample size would certainly be much smaller) of liquid oxygen away from our magnet, energy would have to be expended amounting, in joules, to 75 newtons times a distance of roughly 0.1 meters (since the field strength falls off substantially over a distance of a few centimeters). In order of magnitude, let us say the energy is 10 joules. There are \(2 \cdot 10^{25}\) molecules in 1 kilogram of the liquid, so this is less than \(10^{-24}\) joules per molecule. Just to vaporize 1 kilogram of liquid oxygen requires 50,000 calories, or about \(10^{-20}\) joules per molecule, using 1 calorie = 4.18 joules. (Most of that energy is used in separating the molecules from one another.) Whatever may be happening in liquid oxygen at the molecular level as a result of the magnetic field, it is apparently a very minor affair in terms of energy.

Even a strong magnetic field has hardly any effect on chemical processes, including biochemical. You could put your hand and forearm into our 3 tesla solenoid without experiencing any significant sensation or consequence. It is hard to predict whether your arm would prove to be paramagnetic or diamagnetic, but the force on it would be no more than a fraction of an ounce in any case. Conversely, the presence of someone’s hand close to the sample in Fig. 11.2 would perturb the field and change the force on the sample by no more than a few parts in a million. In whole-body imaging with nuclear magnetic resonance, the body is pervaded by a magnetic field of a few tesla in strength with no physiological effects whatsoever. It appears that the only hazards associated with large-scale, strong, steady magnetic fields arise from metal objects in the vicinity. For example, implants containing metal may heat up, move within the body, or malfunction. And there is also the danger that a loose iron object will be snatched by the fringing field and hurled into the magnet. Be careful what you bring into a magnetic resonance imaging (MRI) room!

In its interaction with matter, the magnetic field plays a role utterly different from that of the electric field. The reason is simple and fundamental. Atoms and molecules are made of electrically charged particles that move with velocities generally small compared with the speed of light. A magnetic field exerts no force at all on a stationary electric charge; on a moving charged particle the force is proportional to \(v^2/c^2\).

---

1 This factor of \(v^2/c^2\) follows from Eq. (5.28). The current \(I\) in that equation involves the velocity of charges, and we are assuming that all velocities here are of the same order of magnitude. Of course, if we have a charged particle moving in a region where the electric field is zero and the magnetic field is nonzero, then the magnetic field dominates. But for general random motions of the charges (both the charges creating the fields and the charges affected by the fields), the magnetic force is smaller by a factor of \(v^2/c^2\).
Said in a sloppier way, in SI units the $1/4\pi \varepsilon_0$ factor in Coulomb’s law for the electric field is large, while the $\mu_0/4\pi$ factor in the Biot–Savart law for the magnetic field is small. Electric forces overwhelmingly dominate the atomic scene. As we have remarked before, magnetism appears, in our world at least, to be a relativistic effect. The story would be different if matter were made of magnetically charged particles. We must explain now what magnetic charge means and what its apparent absence signifies.

11.2 The absence of magnetic “charge”

The magnetic field outside a magnetized rod such as a compass needle looks very much like the electric field outside an electrically polarized rod, a rod that has an excess of positive charge at one end, negative charge at the other (Fig. 11.3). It is conceivable that the magnetic field has sources that are related to it in the same way as electric charge is related to the electric field. Then the north pole of the compass needle would be the location of an excess of one kind of magnetic charge, and the south pole would be the location of an excess of the opposite kind. We might call “north charge” positive and “south charge” negative, with magnetic field directed from positive to negative, a rule like that adopted for electric field and electric charge. Historically, that is how our convention about the positive direction of magnetic field was established.²

What we have called magnetic charge has usually been called magnetic pole strength.

This idea is perfectly sound as far as it goes. It becomes even more plausible when we recall that the fundamental equations of the electromagnetic field are symmetrical in $E$ and $\epsilon B$. Why, then, should we not expect to find symmetry in the sources of the field? With magnetic charge as a possible source of the static magnetic field $B$, we would have $\text{div } B \propto \eta$, where $\eta$ stands for the density of magnetic charge, in complete analogy to the electric charge density $\rho$. Two positive magnetic charges (or north poles) would repel one another, and so on.

The trouble is, that is not the way things are. Nature for some reason has not made use of this opportunity. The world around us appears totally asymmetrical in the sense that we find no magnetic charges at all. No one has yet observed an isolated excess of one kind of magnetic charge – an isolated north pole for example. If such a magnetic monopole existed it could be recognized in several ways. Unlike a magnetic dipole,

² In Chapter 6, remember, we established the positive direction of $B$ by reference to a current direction (direction of motion of positive charge) and a right-hand rule. Now north pole means “north-seeking pole” of the compass needle. We know of no reason why the earth’s magnetic polarity should be one way rather than the other. Franklin’s designation of “positive” electricity had nothing to do with any of this. So the fact that it takes a right-hand rule rather than a left-hand rule to make this all consistent is the purest accident.
Figure 11.3.
(a) Two oppositely charged disks (the electrodes showing in cross section as solid black bars) have an electric field that is the same as that of a polarized rod. That is, if you imagine such a rod to occupy the region within the dashed boundary, its external field would be like that shown. The electric field here was made visible by a multitude of tiny black fibers, suspended in oil, which oriented themselves along the field direction. This elegant method of demonstrating electric field configurations is due to Harold M. Waage, Palmer Physical Laboratory, Princeton University, who kindly prepared the original photograph for this illustration (Waage, 1964).
(b) The magnetic field around a magnetized cylinder, shown by the orientation of small pieces of nickel wire, immersed in glycerine. (This attempt to improve on the traditional iron filings demonstration by an adaptation of Waage’s technique was not very successful – the nickel wires tend to join in long strings that are then pulled in toward the magnet.) Theoretically constructed diagrams of the fields in the two systems are shown later in Fig. 11.22.
it would experience a force if placed in a uniform magnetic field. Thus an elementary particle carrying a magnetic charge would be steadily accelerated in a static magnetic field, as a proton or an electron is steadily accelerated in an electric field. Reaching high energy, it could then be detected by its interaction with matter. A traveling magnetic monopole is a magnetic current; it must be encircled by an electric field, as an electric current is encircled by a magnetic field. With strategies based on these unique properties, physicists have looked for magnetic monopoles in many experiments. The search was renewed when a development in the theory of elementary particles suggested that the universe ought to contain at least a few magnetic monopoles, left over from the “big bang” in which it presumably began. But not one magnetic monopole has yet been detected, and it is now evident that if they exist at all they are exceedingly rare. Of course, the proven existence of even one magnetically charged particle would have profound implications, but it would not alter the fact that in matter as we know it, the only sources of the magnetic field are electric currents. As far as we know,

$$\text{div} \mathbf{B} = 0 \quad \text{(everywhere)}$$

(11.1)

This takes us back to the hypothesis of Ampère, his idea that magnetism in matter is to be accounted for by a multitude of tiny rings of electric current distributed through the substance. We begin by studying the magnetic field created by a single current loop at points relatively far from the loop.

### 11.3 The field of a current loop

A closed conducting loop, not necessarily circular, lies in the \(xy\) plane encircling the origin, as in Fig. 11.4(a). A steady current \(I\) flows around the loop. We are interested in the magnetic field this current creates – not near the loop, but at distant points like \(P_1\) in the figure. We assume that \(r_1\), the distance to \(P_1\), is much larger than any dimension of the loop. To simplify the diagram we have located \(P_1\) in the \(yz\) plane; it will turn out that this is no restriction. This is a good place to use the vector potential. We shall compute first the vector potential \(A\) at the location \(P_1\), that is, \(A(0, y_1, z_1)\). From this it will be obvious what the vector potential is at any other point \((x, y, z)\) far from the loop. Then by taking the curl of \(A\) we can get the magnetic field \(B\).

For a current confined to a wire, Eq. (6.46) gives \(A\) as

$$A(0, y_1, z_1) = \frac{\mu_0 I}{4\pi} \int_{\text{loop}} \frac{dl_2}{r_{12}}. \quad (11.2)$$

When we used this equation in Section 6.4, we were concerned only with the contribution of a small segment of the circuit; now we have to integrate around the entire loop. Consider the variation in the denominator
Figure 11.4.
(a) Calculation of the vector potential \( \mathbf{A} \) at a point far from the current loop. (b) Side view, looking in along the \( x \) axis, showing that \( r_{12} \approx r_1 - y_2 \sin \theta \) if \( r_1 \gg y_2 \). (c) Top view, to show that \( \int_{\text{loop}} y_2 \, dx_2 \) is the area of the loop.

\( r_{12} \) as we go around the loop. If \( P_1 \) is far away, the first-order variation in \( r_{12} \) depends only on the coordinate \( y_2 \) of the segment \( dl_2 \), and not on \( x_2 \). This is true because, from the Pythagorean theorem, the contribution to \( r_{12} \) from \( x_2 \) is of second order, whereas the side view in Fig. 11.4(b) shows the first-order contribution from \( y_1 \). Thus, neglecting quantities proportional to \((x_2/r_{12})^2\), we may treat \( r_{12} \) and \( r_{12}' \), which lie on top of one another in the side view, as equal. And in general, to first order in the ratio (loop dimension/distance to \( P_1 \)), we have

\[
 r_{12} \approx r_1 - y_2 \sin \theta. \tag{11.3}
\]

Look now at the two elements of the path \( dl_2 \) and \( dl'_2 \) shown in Fig. 11.4(a). For these the \( dy_2 \) displacements are equal and opposite, and as we have already pointed out, the \( r_{12} \) distances are equal to first order. To this order then, the \( dy_2 \) contributions to the line integral will cancel,
11.3 The field of a current loop

and this will be true for the whole loop. Hence \( \mathbf{A} \) at \( P_1 \) will not have a \( y \) component. Obviously it will not have a \( z \) component either, for \( dz_2 \) is always zero since the current path itself has nowhere a \( z \) component.

However, \( \mathbf{A} \) at \( P_1 \) will have an \( x \) component. The \( x \) component of the vector potential comes from the \( dx_2 \) part of the path integral:

\[
\mathbf{A}(0, y_1, z_1) = \hat{x} \frac{\mu_0 I}{4\pi} \int \frac{dx_2}{r_{12}}. \tag{11.4}
\]

Without spoiling our first-order approximation, we can turn Eq. (11.3) into

\[
\frac{1}{r_{12}} = \frac{1}{r_1 \left(1 - \frac{y_2}{r_1} \sin \theta\right)} \approx \frac{1}{r_1} \left(1 + \frac{y_2 \sin \theta}{r_1}\right), \tag{11.5}
\]

and using this for the integrand, we have

\[
\mathbf{A}(0, y_1, z_1) = \hat{x} \frac{\mu_0 I}{4\pi r_1} \int \left(1 + \frac{y_2 \sin \theta}{r_1}\right) dx_2. \tag{11.6}
\]

In the integration, \( r_1 \) and \( \theta \) are constants. Obviously \( \int dx_2 \) around the loop vanishes. Now \( \int y_2 \, dx_2 \) around the loop is just the area of the loop, regardless of its shape; see Fig. 11.4(c). So we get finally

\[
\mathbf{A}(0, y_1, z_1) = \hat{x} \frac{\mu_0 I \sin \theta}{4\pi r_1^2} \times \text{(area of loop)}. \tag{11.7}
\]

The intuitive reason why this result is nonzero is that the parts of the loop that are closer to \( P_1 \) give larger contributions to the integral, because they have a smaller \( r_{12} \). There is partial, but not complete, cancelation from corresponding pieces of the loop with the same \( x_2 \) value but opposite \( dx_2 \) values.

Here is a simple but crucial point: since the shape of the loop hasn’t mattered, our restriction on \( P_1 \) to the \( yz \) plane cannot make any essential difference. Therefore we must have in Eq. (11.7) the general result we seek, if only we state it generally: the vector potential of a current loop of any shape, at a distance \( r \) from the loop that is much greater than the size of the loop, is a vector perpendicular to the plane containing \( r \) and the normal to the plane of the loop, of magnitude

\[
A = \frac{\mu_0 I a \sin \theta}{4\pi r^2}, \tag{11.8}
\]

where \( a \) stands for the area of the loop.

This vector potential is symmetrical around the axis of the loop, which implies that the field \( \mathbf{B} \) will be symmetrical also. The explanation is that we are considering regions so far from the loop that the details of the shape of the loop have negligible influence. All loops with the same \( \text{current} \times \text{area} \) product produce the same far field. We call the product \( Ia \)
the magnetic dipole moment of the current loop, and denote it by \( \mathbf{m} \). Its units are \( \text{amp} \cdot \text{m}^2 \). The magnetic dipole moment is a vector, its direction defined to be that of the normal to the loop, or that of the vector \( \mathbf{a} \), the directed area of the region surrounded by the loop:

\[
\mathbf{m} = I \mathbf{a}
\]  \hspace{1cm} (11.9)

As for sign, let us agree that the direction of \( \mathbf{m} \) and the sense of positive current flow in the loop are to be related by a right-hand-screw rule, illustrated in Fig. 11.5. (The dipole moment of the loop in Fig. 11.4(a) points downward, according to this rule.) The vector potential for the field of a magnetic dipole \( \mathbf{m} \) can now be written neatly with vectors:

\[
\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \hat{\mathbf{r}}}{r^2}
\]  \hspace{1cm} (11.10)

where \( \hat{\mathbf{r}} \) is a unit vector in the direction from the loop to the point for which \( \mathbf{A} \) is being computed. You can check that this agrees with our convention about sign. Note that the direction of \( \mathbf{A} \) will always be that of the current in the nearest part of the loop.

Figure 11.6 shows a magnetic dipole located at the origin, with the dipole moment vector \( \mathbf{m} \) pointed in the positive \( z \) direction. To express the vector potential at any point \( (x, y, z) \), we observe that \( r^2 = x^2 + y^2 + z^2 \), and \( \sin \theta = \sqrt{x^2 + y^2} / r \). The magnitude \( A \) of the vector potential at that point is given by

\[
A = \frac{\mu_0 m \sin \theta}{4\pi \frac{r^2}{r^3}} = \frac{\mu_0 m \sqrt{x^2 + y^2}}{4\pi r^3}.
\]  \hspace{1cm} (11.11)

Since \( \mathbf{A} \) is tangent to a horizontal circle around the \( z \) axis, its components are

\[
A_x = A \left( \frac{-y}{\sqrt{x^2 + y^2}} \right) = -\frac{\mu_0 m y}{4\pi r^3},
\]

\[
A_y = A \left( \frac{x}{\sqrt{x^2 + y^2}} \right) = \frac{\mu_0 m x}{4\pi r^3},
\]

\[ A_z = 0. \]  \hspace{1cm} (11.12)

Let’s evaluate \( \mathbf{B} \) for a point in the \( xz \) plane, by finding the components of \( \nabla \times \mathbf{A} \) and then (not before!) setting \( y = 0 \):

\[
B_x = (\nabla \times \mathbf{A})_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = -\frac{\mu_0}{4\pi} \frac{\partial}{\partial z} \frac{mx}{(x^2 + y^2 + z^2)^{3/2}} = \frac{\mu_0}{4\pi} \frac{3mxz}{r^5},
\]

\[
B_y = (\nabla \times \mathbf{A})_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} = \frac{\mu_0}{4\pi} \frac{\partial}{\partial x} \frac{my}{(x^2 + y^2 + z^2)^{3/2}} = \frac{\mu_0}{4\pi} \frac{3myz}{r^5},
\]
11.4 The force on a dipole in an external field

Consider a small circular current loop of radius \( r \), placed in the magnetic field of some other current system, such as a solenoid. In Fig. 11.9, a field

\[
B_z = (\nabla \times \mathbf{A})_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = \frac{\mu_0 m}{4\pi} \left[ \frac{-2x^2 + y^2 + z^2}{(x^2 + y^2 + z^2)^{5/2}} + \frac{x^2 - 2y^2 + z^2}{(x^2 + y^2 + z^2)^{5/2}} \right] = \frac{\mu_0 m(z^2 - r^2)}{4\pi r^5}. \tag{11.13}
\]

In the \( xz \) plane, we have \( y = 0, \sin \theta = x/r, \) and \( \cos \theta = z/r \). The field components at any point in that plane are thus given by

\[
B_x = \frac{\mu_0}{4\pi} \frac{3m \sin \theta \cos \theta}{r^3},

B_y = 0,

B_z = \frac{\mu_0 m(3 \cos^2 \theta - 1)}{4\pi r^3}. \tag{11.14}
\]

Now turn back to Section 10.3, where in Eq. (10.17) we expressed the components in the \( xz \) plane of the field \( \mathbf{E} \) of an electric dipole \( \mathbf{p} \), which was situated exactly like our magnetic dipole \( \mathbf{m} \). The expressions are essentially identical, the only changes being \( p \to m \) and \( 1/\epsilon_0 \to \mu_0 \). We have thus found that the magnetic field of a small current loop has, at remote points, the same form as the electric field of two separated charges. We already know what that field, the electric dipole field, looks like. Figure 11.7 is an attempt to suggest the three-dimensional form of the magnetic field \( \mathbf{B} \) arising from our current loop with dipole moment \( \mathbf{m} \). As in the case of the electric dipole, the field is described somewhat more simply in spherical polar coordinates:

\[
B_r = \frac{\mu_0 m}{2\pi r^3} \cos \theta, \quad B_\theta = \frac{\mu_0 m}{4\pi r^3} \sin \theta, \quad B_\phi = 0. \tag{11.15}
\]

The magnetic field close to a current loop is entirely different from the electric field close to a pair of separated positive and negative charges, as the comparison in Fig. 11.8 shows. Note that between the charges the electric field points down, while inside the current ring the magnetic field points up, although the far fields are alike. This reflects the fact that our magnetic field satisfies \( \nabla \cdot \mathbf{B} = 0 \) everywhere, even inside the source. The magnetic field lines don’t end. By near and far we mean, of course, relative to the size of the current loop or the separation of the charges. If we imagine the current ring shrinking in size, the current meanwhile increasing so that the dipole moment \( m = Ia \) remains constant, we approach the infinitesimal magnetic dipole, the counterpart of the infinitesimal electric dipole described in Chapter 10.
Figure 11.8.
(a) The electric field of a pair of equal and opposite charges. Far away it becomes the field of an electric dipole. (b) The magnetic field of a current ring. Far away it becomes the field of a magnetic dipole.
11.4 The force on a dipole in an external field

\[ B \] is drawn that is generally in the \( z \) direction. It is not a uniform field. Instead, it gets weaker as we proceed in the \( z \) direction; this is evident from the fanning out of the field lines. Let us assume, for simplicity, that the field is symmetric about the \( z \) axis. Then it resembles the field near the upper end of the solenoid in Fig. 11.1. The field represented in Fig. 11.9 does not include the magnetic field of the current ring itself. We want to find the force on the current ring caused by the other field, which we shall call, for want of a better name, the \textit{external field}. The net force on the current ring due to its own field is certainly zero, so we are free to ignore its own field in this discussion.

If you study the situation in Fig. 11.9, you will soon conclude that there is a net force on the current ring. It arises because the external field \( B \) has an outward component \( B_r \) everywhere around the ring. (The vertical component \( B_z \) produces a force in the horizontal plane that simply stretches or compresses the ring – negligibly, assuming the ring is fairly rigid.) Therefore if the current flows in the direction indicated, each element of the loop, \( dl \), must be experiencing a downward force of magnitude \( IB_r \) \( dl \) (see Eq. (6.14)). If \( B_r \) has the same magnitude at all points on the ring, as it must in the symmetrically spreading field assumed, the total downward force will have the magnitude

\[
F = 2\pi rIB_r. \tag{11.16}
\]

Now, \( B_r \) can be directly related to the gradient of \( B_z \). Since \( \text{div } B = 0 \) at all points, the net flux of magnetic field out of any volume is zero. Consider a pancake-like cylinder of radius \( r \) and height \( \Delta z \) (Fig. 11.10). The outward flux from the side is \( 2\pi r(\Delta z)B_r \) and the net outward flux from the end surfaces is

\[
\pi r^2[-B_z(z) + B_z(z + \Delta z)], \tag{11.17}
\]

which to the first order in the small distance \( \Delta z \) is \( \pi r^2(\partial B_z/\partial z)\Delta z \). Setting the total outward flux equal to zero gives

\[
0 = \pi r^2(\partial B_z/\partial z)\Delta z + 2\pi rB_r\Delta z, \quad \text{or} \quad B_r = -\frac{r}{2}\frac{\partial B_z}{\partial z}. \tag{11.18}
\]

As a check on the sign, note that, according to Eq. (11.18), \( B_r \) is positive when \( B_z \) is decreasing upward; a glance at the figure shows that to be correct.

The force on the dipole (with upward taken to be positive) can now be expressed in terms of the gradient of the component \( B_z \) of the external field:

\[
F = -2\pi rl\left(-\frac{r}{2}\frac{\partial B_z}{\partial z}\right) = \pi r^2l\frac{\partial B_z}{\partial z}. \tag{11.19}
\]

In the present case, \( \partial B_z/\partial z \) is negative, so the force is correctly downward. In the factor \( \pi r^2l \) we recognize the magnitude \( m \) of the magnetic

Figure 11.9.
A current ring in an inhomogeneous magnetic field. (The field of the ring itself is not shown.) Because of the radial component of the field, \( B_r \), there is a force on the ring as a whole.

Figure 11.10.
Gauss’s theorem can be used to relate \( B_r \) and \( \partial B_z/\partial z \), leading to Eq. (11.18).
dipole moment of our current ring. So the force on the ring can be expressed very simply in terms of the dipole moment:

\[
F = m \frac{\partial B_z}{\partial z} \tag{11.20}
\]

We haven’t proved it, but you will not be surprised to hear that for small loops of any other shape the force depends only on the current \( \times \) area product, that is, on the dipole moment. The shape doesn’t matter. Of course, we are discussing only loops small enough so that only the first-order variation of the external field, over the span of the loop, is significant.

Our ring in Fig. 11.9 has a magnetic dipole moment \( m \) pointing upward, and the force on it is downward. Obviously, if we could reverse the current in the ring, thereby reversing \( m \), the force would reverse its direction. The situation can be summarized as follows.

- Dipole moment parallel to external field: force acts in direction of increasing field strength.
- Dipole moment antiparallel to external field: force acts in direction of decreasing field strength.
- Uniform external field: zero force.

Quite obviously, this is not the most general situation. The moment \( m \) could be pointing at some odd angle with respect to the field \( B \), and the different components of \( B \) could be varying, spatially, in different ways. Given all the similarities between electric and magnetic dipoles, it is tempting to say that the force on a magnetic dipole should take the same form as the force on an electric dipole, given in Eq. (10.26). That is, the \( x \) component of the force on a magnetic dipole \( m \) should be given by

\[
F_x = m \cdot \nabla B_x \quad \text{(incorrect)}, \tag{11.21}
\]

with corresponding formulas for \( F_y \) and \( F_z \). All three components can be combined into the compact relation,

\[
\mathbf{F} = (\mathbf{m} \cdot \nabla)\mathbf{B} \quad \text{(incorrect)}. \tag{11.22}
\]

You can check in Problem 11.4 that in the above setup with the ring, this force reduces to the force in Eq. (11.20).

However, this argument by analogy is risky, because, although the fields due to electric and magnetic dipoles look the same at large distances, the dipoles themselves look very different up close. One consists of two point charges, the other of a loop of current. The far field is irrelevant when dealing with the force on a dipole. It turns out that, although Eq. (11.22) gives the correct force on a magnetic dipole in many cases, it is not correct in general. The correct expression for the force turns out to be

\[
\mathbf{F} = \nabla (\mathbf{m} \cdot \mathbf{B}) \tag{11.23}
\]
11.4 The force on a dipole in an external field

You can check in Problem 11.4 that in the above setup, this also reduces to the force in Eq. (11.20). At first glance it might seem like Eq. (11.23) comes out of the blue, but there is actually very good motivation for it. We will see in Section 11.6 that the energy of a magnetic dipole in a magnetic field is $-m \cdot B$. (But see Feynman et al. (1977), chap. 15, for a discussion of a subtlety about this energy.) So Eq. (11.23) is the familiar statement that the force equals the negative gradient of the energy.

Under what conditions are the force expressions in Eqs. (11.22) and (11.23) equal? Using the “$\nabla(A \cdot B)$” vector identity in Appendix K, along with the fact that $m$ has no spatial dependence, we find

$$\nabla(m \cdot B) = (m \cdot \nabla)B + m \times (\nabla \times B).$$  \hspace{1cm} (11.24)

Our two expressions for the force are therefore equal if $\nabla \times B = 0$. If we deal only with static setups where $\partial E/\partial t = 0$, then the relevant Maxwell equation reduces to Ampère’s law, $\nabla \times B = \mu_0 J$. So we see that the two expressions for the force agree if the setup involves no currents at the location of the dipole (other than the current in the dipole loop itself). This was the case in the above example. However, Problem 11.4 presents a setup where Eqs. (11.22) and (11.23) yield different forces; the task of that problem is to calculate the force explicitly and show that it agrees with Eq. (11.23).

In Eqs. (11.20) and (11.23) the force is in newtons, with the magnetic field gradient in tesla/meter and the magnetic dipole moment $m$ given by Eq. (11.9): $m = Ia$, where $I$ is in amps and $a$ in m$^2$. There are several equivalent ways to express the units of $m$. From Eq. (11.9) the units are

$$[m] = \text{amp-m}^2. \hspace{1cm} (11.25)$$

But, as you can see from Eq. (11.20), we also have

$$[m] = \frac{\text{newtons}}{\text{tesla/m}} = \frac{\text{newton-m}}{\text{tesla}} = \frac{\text{joules}}{\text{tesla}}. \hspace{1cm} (11.26)$$

Looking back at the three cases summarized on p. 538, we can begin to see what must be happening in the experiments described at the beginning of this chapter. A substance located at the position of the sample in Fig. 11.2 would be attracted into the solenoid if it contained magnetic dipoles parallel to the field $B$ of the coil. It would be pushed out of the solenoid if it contained dipoles pointing in the opposite direction, antiparallel to the field. The force would depend on the gradient of the axial field strength, and would be zero at the midpoint of the solenoid. Also, if the total strength of dipole moments in the sample were proportional to the field strength $B$, then in a given position the force would be proportional to $B$ times $\partial B/\partial z$, and hence to the square of the solenoid current. This is the observed behavior in the case of the diamagnetic

\footnote{This is a sufficient condition, but not necessary. Technically all we need is for $\nabla \times B$ to be parallel to $m$. But if we want the two expressions to be equal for any orientation of $m$, then we need $\nabla \times B = 0$.}
Magnetic fields in matter

and the paramagnetic substances. It looks as if the ferromagnetic samples must have possessed a magnetic moment nearly independent of field strength, but we must set them aside for a special discussion anyway.

How does the application of a magnetic field to a substance evoke in the substance magnetic dipole moments with total strength proportional to the applied field? And why should they be parallel to the field in some substances, and oppositely directed in others? If we can answer these questions, we shall be on the way to understanding the physics of diamagnetism and paramagnetism.

11.5 Electric currents in atoms

We know that an atom consists of a positive nucleus surrounded by negative electrons. To describe it fully we would need the concepts of quantum physics. Fortunately, a simple and easily visualized model of an atom is very helpful for understanding diamagnetism. It is a planetary model with the electrons in orbits around the nucleus, like the model in Bohr’s first quantum theory of the hydrogen atom.

We begin with one electron moving at constant speed on a circular path. Since we are not attempting here to explain atomic structure, we shall not inquire into the reasons why the electron has this particular orbit. We ask only, if it does move in such an orbit, what magnetic effects are to be expected? In Fig. 11.11 we see the electron, visualized as a particle carrying a concentrated electric charge $-e$, moving with speed $v$ on a circular path of radius $r$. In the middle is a positive nuclear charge, making the system electrically neutral. But the nucleus, because of its relatively great mass, moves so slowly that its magnetic effects can be neglected.

At any instant, the electron and the positive charge would appear as an electric dipole, but on the time average the electric dipole moment is zero, producing no steady electric field at a distance. We discussed this point in Section 10.5. The magnetic field of the system, far away, is not zero on the time average. Instead, it is just the field of a current ring. This is because, when considering the time average, it can’t make any difference whether we have all the negative charge gathered into one lump, going around the track, or distributed in bits, as in Fig. 11.11(b), to make a uniform endless procession. The current is the amount of charge that passes a given point on the ring, per second. Since the electron makes $v/2\pi r$ revolutions per second, the current is

$$I = \frac{ev}{2\pi r}.$$  \hspace{1cm} (11.27)

The orbiting electron is equivalent to a ring current of this magnitude with the direction of positive flow opposite to $v$, as shown in Fig. 11.11(c). Its far field is therefore that of a magnetic dipole, of strength

---

**Figure 11.11.**
(a) A model of an atom in which one electron moves at speed $v$ in a circular orbit. (b) Equivalent procession of charge. The average electric current is the same as if the charge $-e$ were divided into small bits, forming a rotating ring of charge. (c) The magnetic moment is the product of current and area.
\[
m = \pi r^2 I = \frac{e vr}{2}. \quad (11.28)
\]

Let us note in passing a simple relation between the magnetic moment \(m\) associated with the electron orbit, and the orbital angular momentum \(L\). The angular momentum is a vector of magnitude \(L = m_e vr\), where \(m_e\) denotes the mass of the electron,\(^4\) and it points downward if the electron is revolving in the sense shown in Fig. 11.11(a). Note that the product \(vr\) occurs in both \(m\) and \(L\). With due regard to direction, we can write

\[
\mathbf{m} = \frac{-e}{2m_e} \mathbf{L} \quad (11.29)
\]

This relation involves nothing but fundamental constants, which should make you suspect that it holds quite generally. Indeed that is the case, although we shall not prove it here. It holds for elliptical orbits, and it holds even for the rosette-like orbits that occur in a central field that is not inverse-square. Remember the important property of any orbit in a central field: angular momentum is a constant of the motion. It follows then, from the general relation expressed by Eq. (11.29) (derived by us only for a special case), that wherever angular momentum is conserved, the magnetic moment also remains constant in magnitude and direction. The factor

\[\frac{-e}{2m_e} \quad \text{or} \quad \frac{\text{magnetic moment}}{\text{angular momentum}} \quad (11.30)\]

is called the orbital magnetomechanical ratio for the electron.\(^5\) The intimate connection between magnetic moment and angular momentum is central to any account of atomic magnetism.

Why don’t we notice the magnetic fields of all the electrons orbiting in all the atoms of every substance? The answer must be that there is a mutual cancelation. In an ordinary lump of matter there must be as many electrons going one way as the other. This is to be expected, for there is nothing to make one sense of rotation intrinsically easier than another, or otherwise to distinguish any unique axial direction. There would have to be something in the structure of the material to single out not merely an axis, but a sense of rotation around that axis!

We may picture a piece of matter, in the absence of any external magnetic field, as containing revolving electrons with their various orbital angular momentum vectors and associated orbital magnetic moments

\(^4\) Our choice of the symbol \(m\) for magnetic moment makes it necessary, in this chapter, to use a different symbol for the electron mass. For angular momentum we choose the symbol \(L\), because \(L\) is traditionally used in atomic physics for orbital angular momentum, which is what we consider here. We shall be dealing with speeds \(v\) much less than \(c\), so the nonrelativistic expression for \(L\) will suffice.

\(^5\) Many people use the term gyromagnetic ratio for this quantity. Some call it the magnetogyric ratio. Whatever the name, it is understood that the magnetic moment is the numerator.
Magnetic fields in matter

distributed evenly over all directions in space. Consider those orbits that happen to have their planes approximately parallel to the xy plane, of which there will be about equal numbers with \( m \) up and \( m \) down. Let's find out what happens to one of these orbits when we switch on an external magnetic field in the \( z \) direction.

We will start by analyzing an electromechanical system that doesn't look much like an atom. In Fig. 11.12 there is an object of mass \( M \) and electric charge \( q \), tethered to a fixed point by a cord of fixed length \( r \). This cord provides the centripetal force that holds the object in its circular orbit. The magnitude of that force \( F_0 \) is given, as we know, by

\[
F_0 = \frac{Mv_0^2}{r}.
\]  

In the initial state, Fig. 11.12(a), there is no external magnetic field. Now, by means of some suitable large solenoid, we begin creating a field \( B \) in the negative \( z \) direction, uniform over the whole region at any given time. While this field is growing at the rate \( dB/dt \), there will be an induced electric field \( E \) all around the path, as indicated in Fig. 11.12(b). To find the magnitude of this field \( E \) we note that the rate of change of flux through the circular path is

\[
\frac{d\Phi}{dt} = \pi r^2 \frac{dB}{dt}.
\]  

This determines the line integral of the electric field, which is really all that matters (we only assume for symmetry and simplicity that it is the same all around the path). Faraday’s law, \( \mathcal{E} = -d\Phi/dt \), gives (ignoring the signs)

\[
\int E \cdot dl = \pi r^2 \frac{dB}{dt}.
\]  

The left-hand side equals \( 2\pi rE \), so we find that

\[
E = \frac{r dB}{2 \frac{dt}{dt}}.
\]  

We have ignored signs so far, but if you apply to Fig. 11.12 your favorite rule for finding the direction of an induced electromotive force, you will see that \( E \) must be in a direction to accelerate the body, if \( q \) is a positive charge. The acceleration along the path, \( dv/dt \), is determined by the force \( qE \):

\[
M \frac{dv}{dt} = qE = \frac{qr dB}{2 \frac{dt}{dt}},
\]  

so that we have a relation between the change in \( v \) and the change in \( B \):

\[
dv = \frac{q r}{2M} dB.
\]
The radius $r$ being fixed by the length of the cord, the factor $qr/2M$ is a constant. Let $\Delta v$ denote the net change in $v$ in the whole process of bringing the field up to the final value $B_1$. Then

$$\Delta v = \int_{v_0}^{v_0+\Delta v} dv = \frac{qr}{2M} \int_0^{B_1} dB = \frac{qrb_1}{2M}. \quad (11.37)$$

Note that the time has dropped out – the final velocity is the same whether the change is made slowly or quickly.

The increased speed of the charge in the final state means an increase in the upward-directed magnetic moment $m$. A negatively charged body would have been decelerated under similar circumstances, which would have decreased its downward moment. In either case, then, the application of the field $B_1$ has brought about a change in magnetic moment opposite to the field. From Eq. (11.28), the magnitude of the change in magnetic moment $\Delta m$ is

$$\Delta m = \frac{q}{2} \Delta v = \frac{q^2 r^2}{4M} B_1. \quad (11.38)$$

Likewise for charges, either positive or negative, revolving in the other direction, the induced change in magnetic moment is opposite to the change in applied magnetic field. Figure 11.13 shows this for a positive charge. It appears that the following relation holds for either sign of charge and either direction of revolution:

$$\Delta m = -\frac{q^2 r^2}{4M} B_1 \quad (11.39)$$

In this example we forced $r$ to be constant by using a cord of fixed length. Let us see how the tension in the cord has changed. We shall
Magnetic fields in matter

assume that \( B_1 \) is small enough so that \( \Delta v \ll v_0 \). In the final state we require a centripetal force of magnitude

\[
F_1 = \frac{M(v_0 + \Delta v)^2}{r} \approx \frac{Mv_0^2}{r} + \frac{2Mv_0\Delta v}{r},
\]

neglecting the term proportional to \((\Delta v)^2\). But now the magnetic field itself provides an inward force on the moving charge, given by \( q(v_0 + \Delta v)B_1 \). Using Eq. (11.37) to express \( qB_1 \) in terms of \( \Delta v \), we find that this extra inward force has the magnitude \((v_0 + \Delta v)(2M\Delta v/r)\) which, to first order in \( \Delta v/v_0 \), is \( 2Mv_0 \Delta v/r \). That is just what is needed, according to Eq. (11.40), to avoid any extra demand on our cord! Hence the tension in the cord remains unchanged at the value \( F_0 \).

This points to a surprising conclusion: our result, Eq. (11.39), must be valid for any kind of tethering force, no matter how it varies with radius. Our cord could be replaced by an elastic spring without affecting the outcome – the radius would still be unchanged in the final state. Or to go at once to a system we are interested in, it could be replaced by the Coulomb attraction of a nucleus for an electron. Or it could be the effective force that acts on one electron in an atom containing many electrons, which has a still different dependence on radius.

Let us apply this to an electron in an atom, substituting the electron mass \( m_e \) for \( M \), and \( e^2 \) for \( q^2 \). Now \( \Delta m \) is the magnetic moment induced by the application of a field \( B_1 \) to the atom. In other words, \( \Delta m/B_1 \) is a magnetic polarizability, defined in the same way as the electric polarizability \( \alpha \) we introduced in Section 10.5. Remember that \( \alpha/4\pi\epsilon_0 \) had the dimensions of volume and turned out to be, in order of magnitude, \( 10^{-30} \text{ m}^3 \), roughly the volume of an atom. By Eq. (11.39), the magnetic polarizability due to one electron in an orbit of radius \( r \) is

\[
\frac{\Delta m}{B_1} = -\frac{e^2r^2}{4m_e}.
\]

Taking the orbit radius \( r \) to be the Bohr radius, \( 0.53 \cdot 10^{-10} \text{ m} \), we find

\[
\frac{\Delta m}{B_1} = -\frac{(1.6 \cdot 10^{-19} \text{ C})^2(0.53 \cdot 10^{-10} \text{ m})^2}{4(9.1 \cdot 10^{-31} \text{ kg})} = -2 \cdot 10^{-29} \frac{\text{ C}^2 \text{ m}^2}{\text{ kg}}.
\]

However, this comparison between \( \Delta m \) and \( B_1 \) isn’t quite a fair one, because magnetic fields contain a somewhat arbitrary factor of \( \mu_0/4\pi \) multiplying the factors of current and distance; see the Biot–Savart law in Eq. (6.49). (An analogous issue arose with the electric polarizability in Section 10.5.) A more reasonable comparison would therefore involve \((\mu_0/4\pi)\Delta m \) and \( B_1 \). Since \( \mu_0/4\pi = 1 \cdot 10^{-7} \text{ kg m/C}^2 \), the numerical value of the ratio is simply modified by a factor of \( 10^{-7} \), and we have

\[
\frac{\mu_0}{4\pi} \frac{\Delta m}{B_1} = -2 \cdot 10^{-36} \text{ m}^3.
\]