We have alluded briefly to the fact that an object's electrical resistance depends on its size and shape, but now we are ready to begin making more mathematical statements about it. As suggested by figure 13, increasing a resistor's cross-sectional area is equivalent to adding more resistors in parallel, which will lead to an overall decrease in resistance. Any real resistor with straight, parallel sides can be sliced up into a large number of pieces, each with cross-sectional area of, say, 1 $\mu m^2$. The number, $N$, of such slices is proportional to the total cross-sectional area of the resistor, and by application of the result of the previous example we therefore find that the resistance of an object is inversely proportional to its cross-sectional area.

An analogous relationship holds for water pipes, which is why high-flow trunk lines have to have large cross-sectional areas. To make lots of water (current) flow through a skinny pipe, we'd need an impractically large pressure (voltage) difference.
A voltmeter is really just an ammeter with an internal resistor, and we use a voltmeter in parallel with the thing that we're trying to measure the voltage difference across. This means that any time we measure the voltage drop across a resistor, we're essentially putting two resistors in parallel. The ammeter inside the voltmeter can be ignored for the purpose of analyzing how current flows in the circuit, since it is essentially just some coiled-up wire with a very low resistance.

Now if we are carrying out this measurement on a resistor that is part of a larger circuit, we have changed the behavior of the circuit through our act of measuring. It is as though we had modified the circuit by replacing the resistance $R$ with the smaller equivalent resistance of $R$ and $R_v$ in parallel. It is for this reason that voltmeters are built with the largest possible internal resistance.

As a numerical example, if we use a voltmeter with an internal resistance of 1 MΩ to measure the voltage drop across a one-ohm resistor, the equivalent resistance is 0.999999 Ω, which is not different enough to make any difference. But if we tried to use the same voltmeter to measure the voltage drop across a 2 MΩ resistor, we would be reducing the resistance of that part of the circuit by a factor of three, which would produce a drastic change in the behavior of the whole circuit.

This is the reason why you can’t use a voltmeter to measure the voltage difference between two different points in mid-air, or between the ends of a piece of wood. This is by no means a stupid thing to want to do, since the world around us is not a constant-voltage environment, the most extreme example being when an electrical storm is brewing. But it will not work with an ordinary voltmeter because the resistance of the air or the wood is many gigaohms. The effect of waving a pair of voltmeter probes around in the air is that we provide a reuniting path for the positive and negative charges that have been separated — through the voltmeter itself, which is a good conductor compared to the air. This reduces to zero the voltage difference we were trying to measure.

In general, a voltmeter that has been set up with an open circuit (or a very large resistance) between its probes is said to be “floating.” An old-fashioned analog voltmeter of the type described here will read zero when left floating, the same as when it was sitting on the shelf. A floating digital voltmeter usually shows an error message.

### Series resistances

The two basic circuit layouts are parallel and series, so a pair of resistors in series, aa/1, is another of the most basic circuits we can make. By conservation of charge, all the current that flows through one resistor must also flow through the other (as well as through the
battery): 

\[ I_1 = I_2. \]

The only way the information about the two resistance values is going to be useful is if we can apply Ohm’s law, which will relate the resistance of each resistor to the current flowing through it and the voltage difference across it. Figure aa/2 shows the three constant-voltage areas. Voltage differences are more physically significant than voltages, so we define symbols for the voltage differences across the two resistors in figure aa/3.

We have three constant-voltage areas, with symbols for the difference in voltage between every possible pair of them. These three voltage differences must be related to each other. It is as though I tell you that Fred is a foot taller than Ginger, Ginger is a foot taller than Sally, and Fred is two feet taller than Sally. The information is redundant, and you really only needed two of the three pieces of data to infer the third. In the case of our voltage differences, we have

\[ |\Delta V_1| + |\Delta V_2| = |\Delta V_{\text{battery}}|. \]

The absolute value signs are because of the ambiguity in how we define our voltage differences. If we reversed the two probes of the voltmeter, we would get a result with the opposite sign. Digital voltmeters will actually provide a minus sign on the screen if the wire connected to the “V” plug is lower in voltage than the one connected to the “COM” plug. Analog voltmeters pin the needle against a peg if you try to use them to measure negative voltages, so you have to fiddle to get the leads connected the right way, and then supply any necessary minus sign yourself.

Figure aa/4 shows a standard way of taking care of the ambiguity in signs. For each of the three voltage measurements around the loop, we keep the same probe (the darker one) on the clockwise side. It is as though the voltmeter was sidling around the circuit like a crab, without ever “crossing its legs.” With this convention, the relationship among the voltage drops becomes

\[ \Delta V_1 + \Delta V_2 = -\Delta V_{\text{battery}}, \]

or, in more symmetrical form,

\[ \Delta V_1 + \Delta V_2 + \Delta V_{\text{battery}} = 0. \]

More generally, this is known as the loop rule for analyzing circuits: the **loop rule**

Assuming the standard convention for plus and minus signs, the sum of the voltage drops around any closed loop in a DC circuit must be zero.

Looking for an exception to the loop rule would be like asking for a hike that would be downhill all the way and that would come back to its starting point!
For the circuit we set out to analyze, the equation

$$\Delta V_1 + \Delta V_2 + \Delta V_{\text{battery}} = 0$$

can now be rewritten by applying Ohm’s law to each resistor:

$$I_1 R_1 + I_2 R_2 + \Delta V_{\text{battery}} = 0.$$ 

The currents are the same, so we can factor them out:

$$I (R_1 + R_2) + \Delta V_{\text{battery}} = 0,$$

and this is the same result we would have gotten if we had been analyzing a one-resistor circuit with resistance $R_1 + R_2$. Thus the equivalent resistance of resistors in series equals the sum of their resistances.

### Two lightbulbs in series example 15

If two identical lightbulbs are placed in series, how do their brightnesses compare with the brightness of a single bulb?

If taken as a whole, the pair of bulbs act like a doubled resistance, so they will draw half as much current from the wall. Each bulb will be dimmer than a single bulb would have been.

The total power dissipated by the circuit is $I \Delta V$. The voltage drop across the whole circuit is the same as before, but the current is halved, so the two-bulb circuit draws half as much total power as the one-bulb circuit. Each bulb draws one-quarter of the normal power.

Roughly speaking, we might expect this to result in one quarter the light being produced by each bulb, but in reality lightbulbs waste quite a high percentage of their power in the form of heat and wavelengths of light that are not visible (infrared and ultraviolet). Less light will be produced, but it’s hard to predict exactly how much less, since the efficiency of the bulbs will be changed by operating them under different conditions.

### More than two equal resistances in series example 16

By straightforward application of the divide-and-conquer technique discussed in the previous section, we find that the equivalent resistance of $N$ identical resistances $R$ in series will be $NR$.

### Dependence of resistance on length example 17

In the previous section, we proved that resistance is inversely proportional to cross-sectional area. By equivalent reason about resistances in series, we find that resistance is proportional to length. Analogously, it is harder to blow through a long straw than through a short one.
Combining the results of examples 13 and 17, we find that the resistance of an object with straight, parallel sides is given by

\[ R = (\text{constant}) \cdot \frac{L}{A} \]

The proportionality constant is called the resistivity, and it depends only on the substance of which the object is made. A resistivity measurement could be used, for instance, to help identify a sample of an unknown substance.

**Choice of high voltage for power lines**

Thomas Edison got involved in a famous technological controversy over the voltage difference that should be used for electrical power lines. At this time, the public was unfamiliar with electricity, and easily scared by it. The president of the United States, for instance, refused to have electrical lighting in the White House when it first became commercially available because he considered it unsafe, preferring the known fire hazard of oil lamps to the mysterious dangers of electricity. Mainly as a way to overcome public fear, Edison believed that power should be transmitted using small voltages, and he publicized his opinion by giving demonstrations at which a dog was lured into position to be killed by a large voltage difference between two sheets of metal on the ground. (Edison’s opponents also advocated alternating current rather than direct current, and AC is more dangerous than DC as well. As we will discuss later, AC can be easily stepped up and down to the desired voltage level using a device called a transformer.)

Now if we want to deliver a certain amount of power \( P_L \) to a load such as an electric lightbulb, we are constrained only by the equation \( P_L = I \Delta V_L \). We can deliver any amount of power we wish, even with a low voltage, if we are willing to use large currents. Modern electrical distribution networks, however, use dangerously high voltage differences of tens of thousands of volts. Why did Edison lose the debate?

It boils down to money. The electric company must deliver the amount of power \( P_L \) desired by the customer through a transmission line whose resistance \( R_T \) is fixed by economics and geography. The same current flows through both the load and the transmission line, dissipating power usefully in the former and wastefully in the latter. The efficiency of the system is

\[
\text{efficiency} = \frac{\text{power paid for by the customer}}{\text{power paid for by the utility}} = \frac{P_L}{P_L + P_T} = \frac{1}{1 + P_T/P_L}
\]
Putting ourselves in the shoes of the electric company, we wish to get rid of the variable $P_T$, since it is something we control only indirectly by our choice of $\Delta V_T$ and $I$. Substituting $P_T = I\Delta V_T$, we find

\[
\text{efficiency} = \frac{1}{1 + \frac{I\Delta V_T}{P_L}}
\]

We assume the transmission line (but not necessarily the load) is ohmic, so substituting $\Delta V_T = IR_T$ gives

\[
\text{efficiency} = \frac{1}{1 + \frac{I^2 R_T}{P_L}}
\]

This quantity can clearly be maximized by making $I$ as small as possible, since we will then be dividing by the smallest possible quantity on the bottom of the fraction. A low-current circuit can only deliver significant amounts of power if it uses high voltages, which is why electrical transmission systems use dangerous high voltages.

---

**Two ways of handling signs**

The figure above shows two ways of visualizing the loop rule and handling the signs involved. In panel 1, each circuit element is labeled with the voltage drop across it.

In 2, the crab is a voltmeter whose reading is the voltage on the white claw minus the voltage on the black claw. The crab can’t flip over. It can only scuttle sideways as it moves around the loop that we’ve chosen, consisting of four resistors. The sum of the four readings is zero.

Panel 3 shows a visualization of the same circuit in which voltage is like height. The stick figure on the ledge wants to get down to...
the ground by doing a series of hops. He has two ways: do the 3 V drop and then the 1 V drop, or do the 2 V and the other 2 V. Here we treat all the voltage differences as positive numbers. This method works nicely if you’re pretty sure for each resistor in the circuit which end is the higher voltage.

Example 20.

A complicated circuit

> All seven resistors in the left-hand panel of figure ab are identical. Initially, the switch S is open as shown in the figure, and the current through resistor A is \( I_0 \). The switch is then closed. Find the current through resistor B, after the switch is closed, in terms of \( I_0 \).

> The second panel shows the circuit redrawn for simplicity, in the initial condition with the switch open. When the switch is open, no current can flow through the central resistor, so we may as well ignore it. I’ve also redrawn the junctions, without changing what’s connected to what. This is the kind of mental rearranging that you’ll eventually learn to do automatically from experience with analyzing circuits. The redrawn version makes it easier to see what’s happening with the current. Charge is conserved, so any charge that flows past point 1 in the circuit must also flow past points 2 and 3. This would have been harder to reason about by applying the junction rule to the original version, which appears to have nine separate junctions.

In the new version, it’s also clear that the circuit has a great deal of symmetry. We could flip over each parallel pair of identical resistors without changing what’s connected to what, so that makes it clear that the voltage drops and currents must be equal for the members of each pair. We can also prove this by using the loop rule. The loop rule says that the two voltage drops in loop 4 must be equal, and similarly for loops 5 and 6. Since the resistors obey Ohm’s law, equal voltage drops across them also imply equal cur-
rents. That means that when the current at point 1 comes to the top junction, exactly half of it goes through each resistor. Then the current reunites at 2, splits between the next pair, and so on. We conclude that each of the six resistors in the circuit experiences the same voltage drop and the same current. Applying the loop rule to loop 7, we find that the sum of the three voltage drops across the three left-hand resistors equals the battery’s voltage, \( V \), so each resistor in the circuit experiences a voltage drop \( V/3 \). Letting \( R \) stand for the resistance of one of the resistors, we find that the current through resistor B, which is the same as the currents through all the others, is given by \( I_0 = V/3R \).

We now pass to the case where the switch is closed, as shown in the third panel. The battery’s voltage is the same as before, and each resistor’s resistance is the same, so we can still use the same symbols \( V \) and \( R \) for them. It is no longer true, however, that each resistor feels a voltage drop \( V/3 \). The equivalent resistance of the whole circuit is \( R/2 + R/3 + R/2 = 4R/3 \), so the total current drawn from the battery is \( 3V/4R \). In the middle group of resistors, this current is split three ways, so the new current through B is \( (1/3)(3V/4R) = V/4R = 3I_0/4 \).

Interpreting this result, we see that it comes from two effects that partially cancel. Closing the switch reduces the equivalent resistance of the circuit by giving charge another way to flow, and increases the amount of current drawn from the battery. Resistor B, however, only gets a 1/3 share of this greater current, not 1/2. The second effect turns out to be bigger than first, and therefore the current through resistor B is lessened over all.

**Getting killed by your ammeter**

As with a voltmeter, an ammeter can give erroneous readings if it is used in such a way that it changes the behavior the circuit. An ammeter is used in series, so if it is used to measure the current through a resistor, the resistor’s value will effectively be changed to \( R + R_a \), where \( R_a \) is the resistance of the ammeter. Ammeters are designed with very low resistances in order to make it unlikely that \( R + R_a \) will be significantly different from \( R \).

In fact, the real hazard is death, not a wrong reading! Virtually the only circuits whose resistances are significantly less than that of an ammeter are those designed to carry huge currents. An ammeter inserted in such a circuit can easily melt. When I was working at a laboratory funded by the Department of Energy, we got periodic bulletins from the DOE safety office about serious accidents at other sites, and they held a certain ghoulish fascination. One of these was about a DOE worker who was completely incinerated by the explosion created when he inserted an ordinary Radio Shack ammeter into a high-current circuit. Later estimates showed that the heat was probably so intense that the explosion
was a ball of plasma — a gas so hot that its atoms have been ionized.

**Discussion question**

A We have stated the loop rule in a symmetric form where a series of voltage drops adds up to zero. To do this, we had to define a standard way of connecting the voltmeter to the circuit so that the plus and minus signs would come out right. Suppose we wish to restate the junction rule in a similar symmetric way, so that instead of equating the current coming in to the current going out, it simply states that a certain sum of currents at a junction adds up to zero. What standard way of inserting the ammeter would we have to use to make this work?
Summary

Selected vocabulary

charge . . . . . . a numerical rating of how strongly an object participates in electrical forces

coulomb (C) . . . the unit of electrical charge

current . . . . . . the rate at which charge crosses a certain boundary

ampere . . . . . . the metric unit of current, one coulomb per second; also “amp”

ammeter . . . . . a device for measuring electrical current

circuit . . . . . . an electrical device in which charge can come back to its starting point and be recycled rather than getting stuck in a dead end

open circuit . . . a circuit that does not function because it has a gap in it

short circuit . . . a circuit that does not function because charge is given a low-resistance “shortcut” path that it can follow, instead of the path that makes it do something useful

voltage . . . . . . electrical potential energy per unit charge that will be possessed by a charged particle at a certain point in space

volt . . . . . . . . the metric unit of voltage, one joule per coulomb

voltmeter . . . . . a device for measuring voltage differences

ohmic . . . . . . . describes a substance in which the flow of current between two points is proportional to the voltage difference between them

resistance . . . . the ratio of the voltage difference to the current in an object made of an ohmic substance

ohm . . . . . . . . the metric unit of electrical resistance, one volt per ampere

Notation

$q$ . . . . . . . . charge

$I$ . . . . . . . . current

$A$ . . . . . . . . units of amperes

$V$ . . . . . . . . voltage

$V$ . . . . . . . . units of volts

$R$ . . . . . . . . resistance

$\Omega$ . . . . . . . units of ohms

Other terminology and notation

electric potential rather than the more informal “voltage” used here; despite the misleading name, it is not the same as electric potential energy

eV . . . . . . . . a unit of energy, equal to $e$ multiplied by 1 volt; $1.6 \times 10^{-19}$ joules
All the forces we encounter in everyday life boil down to two basic types: gravitational forces and electrical forces. A force such as friction or a “sticky force” arises from electrical forces between individual atoms.

Just as we use the word “mass” to describe how strongly an object participates in gravitational forces, we use the word “charge” for the intensity of its electrical forces. There are two types of charge. Two charges of the same type repel each other, but objects whose charges are different attract each other. Charge is measured in units of coulombs (C).

Mobile charged particle model: A great many phenomena are easily understood if we imagine matter as containing two types of charged particles, which are at least partially able to move around.

Positive and negative charge: Ordinary objects that have not been specially prepared have both types of charge spread evenly throughout them in equal amounts. The object will then tend not to exert electrical forces on any other object, since any attraction due to one type of charge will be balanced by an equal repulsion from the other. (We say “tend not to” because bringing the object near an object with unbalanced amounts of charge could cause its charges to separate from each other, and the force would no longer cancel due to the unequal distances.) It therefore makes sense to describe the two types of charge using positive and negative signs, so that an unprepared object will have zero total charge.

The Coulomb force law states that the magnitude of the electrical force between two charged particles is given by \[ |F| = k|q_1||q_2|/r^2. \]

Conservation of charge: An even more fundamental reason for using positive and negative signs for charge is that with this definition the total charge of a closed system is a conserved quantity.

All electrical phenomena are alike in that they arise from the presence or motion of charge. Most practical electrical devices are based on the motion of charge around a complete circuit, so that the charge can be recycled and does not hit any dead ends. The most useful measure of the flow of charge is current, \[ I = \Delta q/\Delta t. \]

An electrical device whose job is to transform energy from one form into another, e.g., a lightbulb, uses power at a rate which depends both on how rapidly charge is flowing through it and on how much work is done on each unit of charge. The latter quantity is known as the voltage difference between the point where the current enters the device and the point where the current leaves it. Since there is a type of potential energy associated with electrical forces, the amount of work they do is equal to the difference in potential energy between the two points, and we therefore define voltage differences directly
in terms of potential energy, $\Delta V = \Delta P_{\text{elec}}/q$. The rate of power dissipation is $P = I\Delta V$.

Many important electrical phenomena can only be explained if we understand the mechanisms of current flow at the atomic level. In metals, currents are carried by electrons, in liquids by ions. Gases are normally poor conductors unless their atoms are subjected to such intense electrical forces that the atoms become ionized.

Many substances, including all solids, respond to electrical forces in such a way that the flow of current between two points is proportional to the voltage difference between those points. Such a substance is called ohmic, and an object made out of an ohmic substance can be rated in terms of its resistance, $R = \Delta V/I$. An important corollary is that a perfect conductor, with $R = 0$, must have constant voltage everywhere within it.

A schematic is a drawing of a circuit that standardizes and stylizes its features to make it easier to understand. Any circuit can be broken down into smaller parts. For instance, one big circuit may be understood as two small circuits in series, another as three circuits in parallel. When circuit elements are combined in parallel and in series, we have two basic rules to guide us in understanding how the parts function as a whole:

**the junction rule:** In any circuit that is not storing or releasing charge, conservation of charge implies that the total current flowing out of any junction must be the same as the total flowing in.

**the loop rule:** Assuming the standard convention for plus and minus signs, the sum of the voltage drops around any closed loop in a circuit must be zero.

The simplest application of these rules is to pairs of resistors combined in series or parallel. In such cases, the pair of resistors acts just like a single unit with a certain resistance value, called their equivalent resistance. Resistances in series add to produce a larger equivalent resistance,

$$R_{\text{series}} = R_1 + R_2,$$

because the current has to fight its way through both resistances. Parallel resistors combine to produce an equivalent resistance that is smaller than either individual resistance,

$$R_{\text{parallel}} = \left( \frac{1}{R_1} + \frac{1}{R_2} \right)^{-1},$$

because the current has two different paths open to it.

An important example of resistances in parallel and series is the use of voltmeters and ammeters in resistive circuits. A voltmeter acts
as a large resistance in parallel with the resistor across which the voltage drop is being measured. The fact that its resistance is not infinite means that it alters the circuit it is being used to investigate, producing a lower equivalent resistance. An ammeter acts as a small resistance in series with the circuit through which the current is to be determined. Its resistance is not quite zero, which leads to an increase in the resistance of the circuit being tested.
Problems

Key
✓ A computerized answer check is available online.
∫ A problem that requires calculus.
* A difficult problem.

1 The figure shows a neuron, which is the type of cell your nerves are made of. Neurons serve to transmit sensory information to the brain, and commands from the brain to the muscles. All this data is transmitted electrically, but even when the cell is resting and not transmitting any information, there is a layer of negative electrical charge on the inside of the cell membrane, and a layer of positive charge just outside it. This charge is in the form of various ions dissolved in the interior and exterior fluids. Why would the negative charge remain plastered against the inside surface of the membrane, and likewise why doesn’t the positive charge wander away from the outside surface?

2 A helium atom finds itself momentarily in this arrangement. Find the direction and magnitude of the force acting on the right-hand electron. The two protons in the nucleus are so close together (∼ 1 fm) that you can consider them as being right on top of each other. As discussed in chapter 26, the charge of an electron is \(-e\), and the charge of a proton \(e\), where \(e = 1.60 \times 10^{-19} \text{ C}\). ✓

3 The helium atom of problem 2 has some new experiences, goes through some life changes, and later on finds itself in the configuration shown here. What are the direction and magnitude of the force acting on the bottom electron? (Draw a sketch to make clear the definition you are using for the angle that gives direction.) ✓

4 Suppose you are holding your hands in front of you, 10 cm apart.
(a) Estimate the total number of electrons in each hand. ✓
(b) Estimate the total repulsive force of all the electrons in one hand on all the electrons in the other. ✓
(c) Why don’t you feel your hands repelling each other?
(d) Estimate how much the charge of a proton could differ in magnitude from the charge of an electron without creating a noticeable force between your hands.
5  As discussed in more detail in section 26.4, a nucleus contains protons, which have positive charge, and neutrons, which have zero charge. If only the electrical force existed, a nucleus would immediately fly apart due to electrical repulsion. However, there is also another force, called the strong nuclear force, which keeps this from happening. Suppose that a proton in a lead nucleus wanders out to the surface of the nucleus, and experiences a strong nuclear force of about 8 kN from the nearby neutrons and protons pulling it back in. Compare this numerically to the repulsive electrical force from the other protons, and verify that the net force is attractive. A lead nucleus is very nearly spherical, is about 6.5 fm in radius, and contains 82 protons, each with a charge of $+e$, where $e = 1.60 \times 10^{-19}$ C.

6  The subatomic particles called muons behave exactly like electrons, except that a muon’s mass is greater by a factor of 206.77. Muons are continually bombarding the Earth as part of the stream of particles from space known as cosmic rays. When a muon strikes an atom, it can displace one of its electrons. If the atom happens to be a hydrogen atom, then the muon takes up an orbit that is on the average 206.77 times closer to the proton than the orbit of the ejected electron. How many times greater is the electric force experienced by the muon than that previously felt by the electron?  

7  The Earth and Moon are bound together by gravity. If, instead, the force of attraction were the result of each having a charge of the same magnitude but opposite in sign, find the quantity of charge that would have to be placed on each to produce the required force.

8  The figure shows one layer of the three-dimensional structure of a salt crystal. The atoms extend much farther off in all directions, but only a six-by-six square is shown here. The larger circles are the chlorine ions, which have charges of $-e$, where $e = 1.60 \times 10^{-19}$ C. The smaller circles are sodium ions, with charges of $+e$. The center-to-center distance between neighboring ions is about 0.3 nm. Real crystals are never perfect, and the crystal shown here has two defects: a missing atom at one location, and an extra lithium atom, shown as a grey circle, inserted in one of the small gaps. If the lithium atom has a charge of $+e$, what is the direction and magnitude of the total force on it? Assume there are no other defects nearby in the crystal besides the two shown here.

$\triangleright$ Hint, p. 1028  

Problem 8.
9 In the semifinals of an electrostatic croquet tournament, Jessica hits her positively charged ball, sending it across the playing field, rolling to the left along the $x$ axis. It is repelled by two other positive charges. These two equal charges are fixed on the $y$ axis at the locations shown in the figure. (a) Express the force on the ball in terms of the ball’s position, $x$. (b) At what value of $x$ does the ball experience the greatest deceleration? Express your answer in terms of $b$. [Based on a problem by Halliday and Resnick.]

10 In a wire carrying a current of 1.0 pA, how long do you have to wait, on the average, for the next electron to pass a given point? Express your answer in units of microseconds. The charge of an electron is $-e = -1.60 \times 10^{-19} \text{ C}$. 

11 Referring back to our old friend the neuron from problem 1 on page 614, let’s now consider what happens when the nerve is stimulated to transmit information. When the blob at the top (the cell body) is stimulated, it causes $\text{Na}^+$ ions to rush into the top of the tail (axon). This electrical pulse will then travel down the axon, like a flame burning down from the end of a fuse, with the $\text{Na}^+$ ions at each point first going out and then coming back in. If $10^{10}$ $\text{Na}^+$ ions cross the cell membrane in 0.5 ms, what amount of current is created? The charge of a $\text{Na}^+$ ion is $+e = 1.60 \times 10^{-19} \text{ C}$. 

12 If a typical light bulb draws about 900 mA from a 110-V household circuit, what is its resistance? (Don’t worry about the fact that it’s alternating current.)

13 A resistor has a voltage difference $\Delta V$ across it, causing a current $I$ to flow.
(a) Find an equation for the power it dissipates as heat in terms of the variables $I$ and $R$ only, eliminating $\Delta V$.
(b) If an electrical line coming to your house is to carry a given amount of current, interpret your equation from part a to explain whether the wire’s resistance should be small, or large.

14 (a) Express the power dissipated by a resistor in terms of $R$ and $\Delta V$ only, eliminating $I$.
(b) Electrical receptacles in your home are mostly 110 V, but circuits for electric stoves, air conditioners, and washers and driers are usually 220 V. The two types of circuits have differently shaped receptacles. Suppose you rewire the plug of a drier so that it can be plugged in to a 110 V receptacle. The resistor that forms the heating element of the drier would normally draw 200 W. How much power does it actually draw now?

15 Use the result of problem 42 on page 622 to find an equation for the voltage at a point in space at a distance $r$ from a point charge $Q$. (Take your $V=0$ distance to be anywhere you like.)
16 You are given a battery, a flashlight bulb, and a single piece of wire. Draw at least two configurations of these items that would result in lighting up the bulb, and at least two that would not light it. (Don’t draw schematics.) If you’re not sure what’s going on, borrow the materials from your instructor and try it. Note that the bulb has two electrical contacts: one is the threaded metal jacket, and the other is the tip (at the bottom in the figure). [Problem by Arnold Arons.]

17 You have to do different things with a circuit to measure current than to measure a voltage difference. Which would be more practical for a printed circuit board, in which the wires are actually strips of metal embedded inside the board? ▶ Solution, p. 1028

18 Problem 18 has been replaced with problem 43.

19 (a) You take an LP record out of its sleeve, and it acquires a static charge of 1 nC. You play it at the normal speed of $33\frac{1}{3}$ r.p.m., and the charge moving in a circle creates an electric current. What is the current, in amperes? \( \sqrt{\text{ }} \) 
(b) Although the planetary model of the atom can be made to work with any value for the radius of the electrons’ orbits, more advanced models that we will study later in this course predict definite radii. If the electron is imagined as circling around the proton at a speed of $2.2 \times 10^6$ m/s, in an orbit with a radius of 0.05 nm, what electric current is created? The charge of an electron is \(-e = -1.60 \times 10^{-19}$ C. \( \sqrt{\text{ }} \)

20 Three charges, each of strength $Q$ ($Q > 0$) form a fixed equilateral triangle with sides of length $b$. You throw a particle of mass $m$ and positive charge $q$ from far away, with an initial speed $v$. Your goal is to get the particle to go to the center of the triangle, your aim is perfect, and you are free to throw from any direction you like. What is the minimum possible value of $v$? You will need the result of problem 42. \( \sqrt{\text{ }} \)

21 Referring back to problem 8, p. 615, about the sodium chloride crystal, suppose the lithium ion is going to jump from the gap it is occupying to one of the four closest neighboring gaps. Which one will it jump to, and if it starts from rest, how fast will it be going by the time it gets there? (It will keep on moving and accelerating after that, but that does not concern us.) You will need the result of problem 42. ▶ Hint, p. 1028 \( \sqrt{\text{ }} \)

22 We have referred to resistors dissipating heat, i.e., we have assumed that $P = I\Delta V$ is always greater than zero. Could $I\Delta V$ come out to be negative for a resistor? If so, could one make a refrigerator by hooking up a resistor in such a way that it absorbed heat instead of dissipating it?
Hybrid and electric cars have been gradually gaining market share, but during the same period of time, manufacturers such as Porsche have also begun designing and selling cars with “mild hybrid” systems, in which power-hungry parts like water pumps are powered by a higher-voltage battery rather than running directly on shafts from the motor. Traditionally, car batteries have been 12 volts. Car companies have dithered over what voltage to use as the standard for mild hybrids, building systems based on 36 V, 42 V, and 48 V. For the purposes of this problem, we consider 36 V.

(a) Suppose the battery in a new car is used to run a device that requires the same amount of power as the corresponding device in the old car. Based on the sample figures above, how would the currents handled by the wires in one of the new cars compare with the currents in the old ones? \[ \sqrt{\text{ }} \]

(b) The real purpose of the greater voltage is to handle devices that need more power. Can you guess why they decided to change to higher-voltage batteries rather than increasing the power without increasing the voltage?

(a) Many battery-operated devices take more than one battery. If you look closely in the battery compartment, you will see that the batteries are wired in series. Consider a flashlight circuit. What does the loop rule tell you about the effect of putting several batteries in series in this way?

(b) The cells of an electric eel’s nervous system are not that different from ours — each cell can develop a voltage difference across it of somewhere on the order of one volt. How, then, do you think an electric eel can create voltages of thousands of volts between different parts of its body?

The heating element of an electric stove is connected in series with a switch that opens and closes many times per second. When you turn the knob up for more power, the fraction of the time that the switch is closed increases. Suppose someone suggests a simpler alternative for controlling the power by putting the heating element in series with a variable resistor controlled by the knob. (With the knob turned all the way clockwise, the variable resistor’s resistance is nearly zero, and when it’s all the way counterclockwise, its resistance is essentially infinite.) (a) Draw schematics. (b) Why would the simpler design be undesirable?
26. A 1.0 Ω toaster and a 2.0 Ω lamp are connected in parallel with the 110-V supply of your house. (Ignore the fact that the voltage is AC rather than DC.)
(a) Draw a schematic of the circuit.
(b) For each of the three components in the circuit, find the current passing through it and the voltage drop across it.
(c) Suppose they were instead hooked up in series. Draw a schematic and calculate the same things.

27. Wire is sold in a series of standard diameters, called “gauges.” The difference in diameter between one gauge and the next in the series is about 20%. How would the resistance of a given length of wire compare with the resistance of the same length of wire in the next gauge in the series?

28. The figure shows two possible ways of wiring a flashlight with a switch. Both will serve to turn the bulb on and off, although the switch functions in the opposite sense. Why is method 1 preferable?

29. In the figure, the battery is 9 V.
(a) What are the voltage differences across each light bulb?
(b) What current flows through each of the three components of the circuit?
(c) If a new wire is added to connect points A and B, how will the appearances of the bulbs change? What will be the new voltages and currents?
(d) Suppose no wire is connected from A to B, but the two bulbs are switched. How will the results compare with the results from the original setup as drawn?

30. You have a circuit consisting of two unknown resistors in series, and a second circuit consisting of two unknown resistors in parallel.
(a) What, if anything, would you learn about the resistors in the series circuit by finding that the currents through them were equal?
(b) What if you found out the voltage differences across the resistors in the series circuit were equal?
(c) What would you learn about the resistors in the parallel circuit from knowing that the currents were equal?
(d) What if the voltages in the parallel circuit were equal?
Problem 31. A student in a biology lab is given the following instructions: “Connect the cerebral eraser (C.E.) and the neural depolarizer (N.D.) in parallel with the power supply (P.S.). (Under no circumstances should you ever allow the cerebral eraser to come within 20 cm of your head.) Connect a voltmeter to measure the voltage across the cerebral eraser, and also insert an ammeter in the circuit so that you can make sure you don’t put more than 100 mA through the neural depolarizer.” The diagrams show two lab groups’ attempts to follow the instructions. (a) Translate diagram a into a standard-style schematic. What is correct and incorrect about this group’s setup? (b) Do the same for diagram b.

Problem 31.

(a)  

(b)  

Problem 31.

32. How many different resistance values can be created by combining three unequal resistors? (Don’t count possibilities where not all the resistors are used.)

33. A person in a rural area who has no electricity runs an extremely long extension cord to a friend’s house down the road so she can run an electric light. The cord is so long that its resistance, \( x \), is not negligible. Show that the lamp’s brightness is greatest if its resistance, \( y \), is equal to \( x \). Explain physically why the lamp is dim for values of \( y \) that are too small or too large.

34. What resistance values can be created by combining a 1 k\( \Omega \) resistor and a 10 k\( \Omega \) resistor?  

\( \text{Solution, p. 1028} \)

35. Suppose six identical resistors, each with resistance \( R \), are connected so that they form the edges of a tetrahedron (a pyramid with three sides in addition to the base, i.e., one less side than an Egyptian pyramid). What resistance value or values can be obtained by making connections onto any two points on this arrangement?  

\( \text{Solution, p. 1028} \)
36  The figure shows a circuit containing five lightbulbs connected to a battery. Suppose you’re going to connect one probe of a voltmeter to the circuit at the point marked with a dot. How many unique, nonzero voltage differences could you measure by connecting the other probe to other wires in the circuit?

37  The lightbulbs in the figure are all identical. If you were inserting an ammeter at various places in the circuit, how many unique currents could you measure? If you know that the current measurement will give the same number in more than one place, only count that as one unique current.

38  The bulbs are all identical. Which one doesn’t light up?  ⋆

39  Each bulb has a resistance of one ohm. How much power is drawn from the one-volt battery?  √ ⋆

40  The bulbs all have unequal resistances. Given the three currents shown in the figure, find the currents through bulbs A, B, C, and D.

41  A silk thread is uniformly charged by rubbing it with llama fur. The thread is then dangled vertically above a metal plate and released. As each part of the thread makes contact with the conducting plate, its charge is deposited onto the plate. Since the thread is accelerating due to gravity, the rate of charge deposition increases with time, and by time \( t \) the cumulative amount of charge is \( q = ct^2 \), where \( c \) is a constant. (a) Find the current flowing onto the plate.  √ (b) Suppose that the charge is immediately carried away through a resistance \( R \). Find the power dissipated as heat.  √
(a) Recall from example 7 on p. 346 that the gravitational energy of two gravitationally interacting spheres is given by \( PE = -\frac{Gm_1 m_2}{r} \), where \( r \) is the center-to-center distance. Sketch a graph of \( PE \) as a function of \( r \), making sure that your graph behaves properly at small values of \( r \), where you’re dividing by a small number, and at large ones, where you’re dividing by a large one. Check that your graph behaves properly when a rock is dropped from a larger \( r \) to a smaller one; the rock should lose potential energy as it gains kinetic energy.

(b) Electrical forces are closely analogous to gravitational ones, since both depend on \( 1/r^2 \). Since the forces are analogous, the potential energies should also behave analogously. Using this analogy, write down the expression for the electrical potential energy of two interacting charged particles. The main uncertainty here is the sign out in front. Like masses attract, but like charges repel. To figure out whether you have the right sign in your equation, sketch graphs in the case where both charges are positive, and also in the case where one is positive and one negative; make sure that in both cases, when the charges are released near one another, their motion causes them to lose PE while gaining KE.

In example 8 on p. 592, suppose that the larger sphere has radius \( a \), the smaller one \( b \). (a) Use the result of problem 15 to show that the ratio of the charges on the two spheres is \( q_a/q_b = a/b \). (b) Show that the density of charge (charge per unit area) is the other way around: the charge density on the smaller sphere is greater than that on the larger sphere in the ratio \( a/b \).

Find the current drawn from the battery.

It’s fairly common in electrical circuits for additional, undesirable resistances to occur because of factors such as dirty, corroded, or loose connections. Suppose that a device with resistance \( R \) normally dissipates power \( P \), but due to an additional series resistance \( r \) the total power is reduced to \( P' \). We might, for example, detect this change because the battery powering our device ran down more quickly than normal.

(a) Find the unknown resistance \( r \).

(b) Check that the units of your result make sense.

(c) Check that your result makes sense in the special cases \( P' = P \) and \( P' = 0 \).

(d) Suppose we redefine \( P' \) as the useful power dissipated in \( R \). For example, this would be the change we would notice because a flashlight was dimmer. Find \( r \).
Exercise 21A: Electrical measurements

1. How many different currents could you measure in this circuit? Make a prediction, and then try it.

What do you notice? How does this make sense in terms of the roller coaster metaphor introduced in discussion question 21.5A on p. 587?

What is being used up in the resistor?

2. By connecting probes to these points, how many ways could you measure a voltage? How many of them would be different numbers? Make a prediction, and then do it.

What do you notice? Interpret this using the roller coaster metaphor, and color in parts of the circuit that represent constant voltages.

3. The resistors are unequal. How many different voltages and currents can you measure? Make a prediction, and then try it.

What do you notice? Interpret this using the roller coaster metaphor, and color in parts of the circuit that represent constant voltages.
Exercise 21B: Voltage and current

This exercise is based on one created by Virginia Roundy.

Apparatus:
- DC power supply
- 1.5 volt batteries
- lightbulbs and holders
- wire
- highlighting pens, 3 colors

When you first glance at this exercise, it may look scary and intimidating — all those circuits! However, all those wild-looking circuits can be analyzed using the following four guides to thinking:

1. A circuit has to be complete, i.e., it must be possible for charge to get recycled as it goes around the circuit. If it’s not complete, then charge will build up at a dead end. This built-up charge will repel any other charge that tries to get in, and everything will rapidly grind to a stop.

2. There is constant voltage everywhere along a piece of wire. To apply this rule during this lab, I suggest you use the colored highlighting pens to mark the circuit. For instance, if there’s one whole piece of the circuit that’s all at the same voltage, you could highlight it in yellow. A second piece of the circuit, at some other voltage, could be highlighted in blue.

3. Charge is conserved, so charge can’t “get used up.”

4. You can draw a rollercoaster diagram, like the one shown below. On this kind of diagram, height corresponds to voltage — that’s why the wires are drawn as horizontal tracks.

A Bulb and a Switch

Look at circuit 1, and try to predict what will happen when the switch is open, and what will happen when it’s closed. Write both your predictions in the table on the following page before you build the circuit. When you build the circuit, you don’t need an actual switch like a light switch; just connect and disconnect the banana plugs. Use one of the 1.5 volt batteries as your voltage source.
Did it work the way you expected? If not, try to figure it out with the benefit of hindsight, and write your explanation in the table above.

Exercise 21B: Voltage and current
### Circuit 3

<table>
<thead>
<tr>
<th>switch open</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>switch closed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>observation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Circuit 4

<table>
<thead>
<tr>
<th>switch open</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>switch closed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>explanation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>observation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(If different)
Two Bulbs

Instead of a battery, use the DC power supply, set to 2.4 volts, for circuits 5 and 6. Analyze this one both by highlighting and by drawing a rollercoaster diagram.

![Circuit Diagram](image)

### Circuit 5

<table>
<thead>
<tr>
<th>bulb a</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb a</th>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb b</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb b</th>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Circuit 6

<table>
<thead>
<tr>
<th>bulb a</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb a</th>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb b</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bulb b</th>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Two Batteries

Use batteries for circuits 7-9. Circuits 7 and 8 are both good candidates for rollercoaster diagrams.

Circuit 7

<table>
<thead>
<tr>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
</table>

Circuit 8

<table>
<thead>
<tr>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
</table>

A Final Challenge

Circuit 9

<table>
<thead>
<tr>
<th>bulb a</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>bulb b</th>
<th>prediction</th>
<th>explanation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>observation</th>
<th>explanation (if different)</th>
</tr>
</thead>
</table>
Exercise 21C: Reasoning about circuits

The questions in this exercise can all be solved using some combination of the following approaches:

a) There is constant voltage throughout any conductor.
b) Ohm’s law can be applied to any part of a circuit.
c) Apply the loop rule.
d) Apply the junction rule.

In each case, discuss the question, decide what you think is the right answer, and then try the experiment.

If you’ve already done exercise 21B, skip number 1.

1. The series circuit is changed as shown.

Which reasoning is correct?

- Each bulb now has its sides connected to the two terminals of the battery, so each now has 2.4 V across it instead of 1.2 V. They get brighter.

- Just as in the original circuit, the current goes through one bulb, then the other. It’s just that now the current goes in a figure-8 pattern. The bulbs glow the same as before.

2. A wire is added as shown to the original circuit.

What is wrong with the following reasoning?

The top right bulb will go out, because its two sides are now connected with wire, so there will be no voltage difference across it. The other three bulbs will not be affected.
3. A wire is added as shown to the original circuit.

![Original Circuit Diagram]

What is wrong with the following reasoning?

The current flows out of the right side of the battery. When it hits the first junction, some of it will go left and some will keep going up. The part that goes up lights the top right bulb. The part that turns left then follows the path of least resistance, going through the new wire instead of the bottom bulb. The top bulb stays lit, the bottom one goes out, and others stay the same.

4. What happens when one bulb is unscrewed, leaving an air gap?

![Diagram with One Bulb Unscrewed]

5. This part is optional. You can do it if you finished early and would like an extra challenge.

Predict the voltage drop across each of the three bulbs in part 4, and also predict how the three currents will compare with one another. (You can’t predict the currents in units of amperes, since you don’t know the resistances of the bulbs.) Test your predictions. If your predictions are wrong, try to figure out what’s going on.
Chapter 22
The Nonmechanical Universe

“Okay. Your duties are as follows: Get Breen. I don’t care how you get him, but get him soon. That faker! He posed for twenty years as a scientist without ever being apprehended. Well, I’m going to do some apprehending that’ll make all previous apprehending look like no apprehension at all. You with me?”

“Yes,” said Battle, very much confused. “What’s that thing you have?”

“Piggy-back heat-ray. You transpose the air in its path into an unstable isotope which tends to carry all energy as heat. Then you shoot your juice light, or whatever along the isotopic path and you burn whatever’s on the receiving end. You want a few?”

“No,” said Battle. “I have my gats. What else have you got for offense and defense?” Underbottam opened a cabinet and proudly waved an arm. “Everything,” he said.

“Disintegraters, heat-rays, bombs of every type. And impenetrable shields of energy, massive and portable. What more do I need?”

Cutting-edge science readily infiltrates popular culture, though sometimes in garbled form. The Newtonian imagination populated the universe mostly with that nice solid stuff called matter, which was made of little hard balls called atoms. In the early twentieth century, consumers of pulp fiction and popularized science began to hear of a new image of the universe, full of x-rays, N-rays, and Hertzian waves. What they were beginning to soak up through their skins was a drastic revision of Newton’s concept of a universe made of chunks of matter which happened to interact via forces. In the newly emerging picture, the universe was made of force, or, to be more technically accurate, of ripples in universal fields of force. Unlike the average reader of Cosmic Stories in 1941, you now have enough technical background to understand what a “force field” really is.

22.1 The stage and the actors

Newton’s instantaneous action at a distance

The Newtonian picture has particles interacting with each other by exerting forces from a distance, and these forces are imagined to occur without any time delay. For example, suppose that super-powerful aliens, angered when they hear disco music in our AM radio transmissions, come to our solar system on a mission to cleanse the universe of our aesthetic contamination. They apply a force to our sun, causing it to go flying out of the solar system at a gazillion miles an hour. According to Newton’s laws, the gravitational force of the sun on the earth will immediately start dropping off. This will be detectable on earth, and since sunlight takes eight minutes to get from the sun to the earth, the change in gravitational force will, according to Newton, be the first way in which earthlings learn the bad news — the sun will not visibly start receding until a little later. Although this scenario is fanciful enough to be at home in the pages of Cosmic Stories, it shows a real feature of Newton’s laws: that information can be transmitted from one place in the universe to another with zero time delay, so that transmission and reception occur at exactly the same instant.

Newton was sharp enough to realize that this required a nontrivial assumption, which was that there was some completely objective and well-defined way of saying whether two things happened at exactly the same instant. He stated this assumption explicitly: “Absolute, true, and mathematical time, of itself, and from its own nature flows at a constant rate without regard to anything external...”

No absolute time

Ever since Einstein, we’ve known that this assumption was false. When Einstein first began to develop the theory of relativity, around 1905, the only real-world observations he could draw on were ambiguous and indirect. Today, the evidence is part of everyday life. For example, every time you use a GPS receiver, figure a, you’re using Einstein’s theory of relativity. Somewhere between 1905 and today, technology became good enough to allow conceptually simple experiments that students in the early 20th century could only discuss in terms like “Imagine that we could...” A good jumping-on point is 1971. In that year, J.C. Hafele and R.E. Keating brought atomic clocks aboard commercial airliners, b, and flew around the world, once from east to west and once from west to east. Hafele and Keating observed that there was a discrepancy between the times measured by the traveling clocks and the times measured by similar clocks that stayed home at the U.S. Naval Observatory in Washington. The east-going clock lost time, ending up off by $-59 \pm 10$ nanoseconds, while the west-going one gained $273 \pm 7$ ns. Although this example is particularly dramatic, a large number of other ex-
experiments have also confirmed that time is not absolute, as Newton had imagined.

Nevertheless, the effects that Hafele and Keating observed were small. This makes sense: Newton's laws have already been thoroughly tested by experiments under a wide variety of conditions, so a new theory like relativity must agree with Newton's to a good approximation, within the Newtonian theory's realm of applicability. This requirement of backward-compatibility is known as the correspondence principle.

Causality

It's also reassuring that the effects on time were small compared to the three-day lengths of the plane trips. There was therefore no opportunity for paradoxical scenarios such as one in which the east-going experimenter arrived back in Washington before he left and then convinced himself not to take the trip. A theory that maintains this kind of orderly relationship between cause and effect is said to satisfy causality.

Causality is like a water-hungry front-yard lawn in Los Angeles: we know we want it, but it's not easy to explain why. Even in plain old Newtonian physics, there is no clear distinction between past and future. In figure c, number 18 throws the football to number 25, and the ball obeys Newton's laws of motion. If we took a video of the pass and played it backward, we would see the ball flying from 25 to 18, and Newton's laws would still be satisfied. Nevertheless, we have a strong psychological impression that there is a forward arrow of time. I can remember what the stock market did last year, but I can't remember what it will do next year. Joan of Arc's military victories against England caused the English to burn her at the stake; it's hard to accept that Newton's laws provide an equally good description of a process in which her execution in 1431 caused her to win a battle in 1429. There is no consensus at this point among physicists on the origin and significance of time's arrow, and for our present purposes we don't need to solve this mystery. Instead, we merely note the empirical fact that, regardless of what causality really means and where it really comes from, its behavior is consistent. Specifically, experiments show that if an observer in a certain frame of reference observes that event A causes event B, then observers in other frames agree that A causes B, not the other way around. This is merely a generalization about a large body of experimental results, not a logically necessary assumption. If Keating had gone around the world and arrived back in Washington before he left, it would have disproved this statement about causality.

Time delays in forces exerted at a distance

Relativity is closely related to electricity and magnetism, and we will go into relativity in more detail in chapters 23-27. What we
care about for now is that relativity forbids Newton’s instantaneous action at a distance. For suppose that instantaneous action at a distance existed. It would then be possible to send signals from one place in the universe to another without any time lag. This would allow perfect synchronization of all clocks. But the Hafele-Keating experiment demonstrates that clocks A and B that have been initially synchronized will drift out of sync if one is in motion relative to the other. With instantaneous transmission of signals, we could determine, without having to wait for A and B to be reunited, which was ahead and which was behind. Since they don’t need to be reunited, neither one needs to undergo any acceleration; each clock can fix an inertial frame of reference, with a velocity vector that changes neither its direction nor its magnitude. But this violates the principle that constant-velocity motion is relative, because each clock can be considered to be at rest, in its own frame of reference. Since no experiment has ever detected any violation of the relativity of motion, we conclude that instantaneous action at a distance is impossible.

Since forces can’t be transmitted instantaneously, it becomes natural to imagine force-effects spreading outward from their source like ripples on a pond, and we then have no choice but to impute some physical reality to these ripples. We call them fields, and they have their own independent existence. Chapters 22-24 are mainly about the electric and magnetic fields, although we’ll also talk about the gravitational field. Ripples of the electric and magnetic fields turn out to be light waves. Fields don’t have to wiggle; they can hold still as well. The earth’s magnetic field, for example, is nearly constant, which is why we can use it for direction-finding.

Even empty space, then, is not perfectly featureless. It has measurable properties. For example, we can drop a rock in order to measure the direction of the gravitational field, or use a magnetic compass to find the direction of the magnetic field. This concept made a deep impression on Einstein as a child. He recalled that when he was five years old, the gift of a magnetic compass convinced him that there was “something behind things, something deeply hidden.”

More evidence that fields of force are real: they carry energy.

The smoking-gun argument for this strange notion of traveling force ripples comes from the fact that they carry energy. In figure d/1, Alice and Betty hold positive charges A and B at some distance from one another. If Alice chooses to move her charge closer to Betty’s, d/2, Alice will have to do some mechanical work against the electrical repulsion, burning off some of the calories from that chocolate cheesecake she had at lunch. This reduction in her body’s chemical energy is offset by a corresponding increase in the electrical potential

---

2 As discussed in ch. 24, there are actually two different effects here, one due to motion and one due to gravity.
Fields carry energy. Not only that, but Alice feels the resistance stiffen as the charges get closer together and the repulsion strengthens. She has to do a little extra work, but this is all properly accounted for in the electrical potential energy.

But now suppose, \( d/3 \), that Betty decides to play a trick on Alice by tossing charge B far away just as Alice is getting ready to move charge A. We have already established that Alice can’t feel charge B’s motion instantaneously, so the electric forces must actually be propagated by an electric field. Of course this experiment is utterly impractical, but suppose for the sake of argument that the time it takes the change in the electric field to propagate across the diagram is long enough so that Alice can complete her motion before she feels the effect of B’s disappearance. She is still getting stale information about B’s position. As she moves A to the right, she feels a repulsion, because the field in her region of space is still the field caused by B in its old position. She has burned some chocolate cheesecake calories, and it appears that conservation of energy has been violated, because these calories can’t be properly accounted for by any interaction with B, which is long gone.

If we hope to preserve the law of conservation of energy, then the only possible conclusion is that the electric field itself carries away the cheesecake energy. In fact, this example represents an impractical method of transmitting radio waves. Alice does work on charge A, and that energy goes into the radio waves. Even if B had never existed, the radio waves would still have carried energy, and Alice would still have had to do work in order to create them.

**Discussion questions**

Amy and Bill are flying on spaceships in opposite directions at such high velocities that the relativistic effect on time’s rate of flow is easily noticeable. Motion is relative, so Amy considers herself to be at rest and Bill to be in motion. She says that time is flowing normally for her, but Bill is slow. But Bill can say exactly the same thing. How can they both think the other is slow? Can they settle the disagreement by getting on the radio and seeing whose voice is normal and whose sounds slowed down and Darth-Vadery?
22.2 The gravitational field

Given that fields of force are real, how do we define, measure, and calculate them? A fruitful metaphor will be the wind patterns experienced by a sailing ship. Wherever the ship goes, it will feel a certain amount of force from the wind, and that force will be in a certain direction. The weather is ever-changing, of course, but for now let’s just imagine steady wind patterns. Definitions in physics are operational, i.e., they describe how to measure the thing being defined. The ship’s captain can measure the wind’s “field of force” by going to the location of interest and determining both the direction of the wind and the strength with which it is blowing. Charting all these measurements on a map leads to a depiction of the field of wind force like the one shown in the figure. This is known as the “sea of arrows” method of visualizing a field.

Now let’s see how these concepts are applied to the fundamental force fields of the universe. We’ll start with the gravitational field, which is the easiest to understand. As with the wind patterns, we’ll start by imagining gravity as a static field, even though the existence of the tides proves that there are continual changes in the gravity field in our region of space. Defining the direction of the gravitational field is easy enough: we simply go to the location of interest and measure the direction of the gravitational force on an object, such as a weight tied to the end of a string.

But how should we define the strength of the gravitational field? Gravitational forces are weaker on the moon than on the earth, but we cannot specify the strength of gravity simply by giving a certain number of newtons. The number of newtons of gravitational force depends not just on the strength of the local gravitational field but also on the mass of the object on which we’re testing gravity, our “test mass.” A boulder on the moon feels a stronger gravitational force than a pebble on the earth. We can get around this problem by defining the strength of the gravitational field as the force acting on an object, divided by the object’s mass.

**definition of the gravitational field**

The gravitational field vector, \( \mathbf{g} \), at any location in space is found by placing a test mass \( m_t \) at that point. The field vector is then given by \( \mathbf{g} = \mathbf{F}/m_t \), where \( \mathbf{F} \) is the gravitational force on the test mass.

The magnitude of the gravitational field near the surface of the earth is about 9.8 N/kg, and it’s no coincidence that this number looks familiar, or that the symbol \( g \) is the same as the one for gravitational acceleration. The force of gravity on a test mass will equal \( m_t \mathbf{g} \), where \( \mathbf{g} \) is the gravitational acceleration. Dividing by \( m_t \) simply gives the gravitational acceleration. Why define a new name and new units for the same old quantity? The main reason is...
that it prepares us with the right approach for defining other fields.

The most subtle point about all this is that the gravitational field tells us about what forces would be exerted on a test mass by the earth, sun, moon, and the rest of the universe, if we inserted a test mass at the point in question. The field still exists at all the places where we didn’t measure it.

Gravitational field of the earth
example 1
▷ What is the magnitude of the earth’s gravitational field, in terms of its mass, \( M \), and the distance \( r \) from its center?

▷ Substituting \( |F| = \frac{GMm}{r^2} \) into the definition of the gravitational field, we find \( |g| = \frac{GM}{r^2} \). This expression could be used for the field of any spherically symmetric mass distribution, since the equation we assumed for the gravitational force would apply in any such case.

Sources and sinks

If we make a sea-of-arrows picture of the gravitational fields surrounding the earth, \( f \), the result is evocative of water going down a drain. For this reason, anything that creates an inward-pointing field around itself is called a sink. The earth is a gravitational sink. The term “source” can refer specifically to things that make outward fields, or it can be used as a more general term for both “outies” and “innies.” However confusing the terminology, we know that gravitational fields are only attractive, so we will never find a region of space with an outward-pointing field pattern.

Knowledge of the field is interchangeable with knowledge of its sources (at least in the case of a static, unchanging field). If aliens saw the earth’s gravitational field pattern they could immediately infer the existence of the planet, and conversely if they knew the mass of the earth they could predict its influence on the surrounding gravitational field.

Superposition of fields

A very important fact about all fields of force is that when there is more than one source (or sink), the fields add according to the rules of vector addition. The gravitational field certainly will have this property, since it is defined in terms of the force on a test mass, and forces add like vectors. Superposition is an important characteristic of waves, so the superposition property of fields is consistent with the idea that disturbances can propagate outward as waves in a field.

Reduction in gravity on Io due to Jupiter’s gravity
example 2
▷ The average gravitational field on Jupiter’s moon Io is 1.81 N/kg. By how much is this reduced when Jupiter is directly overhead? Io’s orbit has a radius of \( 4.22 \times 10^8 \) m, and Jupiter’s mass is \( 1.899 \times 10^{27} \) kg.
By the shell theorem, we can treat the Jupiter as if its mass was all concentrated at its center, and likewise for Io. If we visit Io and land at the point where Jupiter is overhead, we are on the same line as these two centers, so the whole problem can be treated one-dimensionally, and vector addition is just like scalar addition. Let’s use positive numbers for downward fields (toward the center of Io) and negative for upward ones. Plugging the appropriate data into the expression derived in example 1, we find that the Jupiter’s contribution to the field is \(-0.71 \text{ N/kg}\). Superposition says that we can find the actual gravitational field by adding up the fields created by Io and Jupiter: \(1.81 - 0.71 \text{ N/kg} = 1.1 \text{ N/kg}\). You might think that this reduction would create some spectacular effects, and make Io an exciting tourist destination. Actually you would not detect any difference if you flew from one side of Io to the other. This is because your body and Io both experience Jupiter’s gravity, so you follow the same orbital curve through the space around Jupiter.

Gravitational waves

Looking back at the argument given on p. 634 for the existence of energy-bearing ripples in the electric field, we see that nowhere was it necessary to appeal to any specific properties of the electrical interaction. We therefore expect energy-carrying gravitational waves to exist, and Einstein’s general theory of relativity does describe such waves and their properties.

![The part of the LIGO gravity wave detector at Hanford Nuclear Reservation, near Richland, Washington. The other half of the detector is in Louisiana.](image)

A Caltech-MIT collaboration has built a pair of gravitational wave detectors called LIGO to search for direct evidence of gravitational
waves. Since they are essentially the most sensitive vibration detectors ever made, they are located in quiet rural areas, and signals are compared between them to make sure that they were not due to passing trucks. The signature of a gravitational wave is if the same wiggle is seen in both detectors within a short time. The detectors are able to sense a vibration that causes a change of $10^{-18}$ m in the distance between the mirrors at the ends of the 4-km vacuum tunnels. This is a thousand times less than the size of an atomic nucleus! In 2016, the collaboration announced the first detection of a gravitational wave, which is believed to have originated from the collision of two black holes. Propagation of gravitational waves at $c$ was verified through multiple methods both by study of the 2016 event and through an event in 2017, interpreted as a collision of two neutron stars, in which both gravitational waves and electromagnetic waves were detected simultaneously.

### 22.3 The electric field

**Definition**

The definition of the electric field is directly analogous to, and has the same motivation as, the definition of the gravitational field:

**definition of the electric field**

The electric field vector, $\mathbf{E}$, at any location in space is found by placing a test charge $q_t$ at that point. The electric field vector is then given by $\mathbf{E} = \mathbf{F}/q_t$, where $\mathbf{F}$ is the electric force on the test charge.

Charges are what create electric fields. Unlike gravity, which is always attractive, electricity displays both attraction and repulsion. A positive charge is a source of electric fields, and a negative one is a sink.

The most difficult point about the definition of the electric field is that the force on a negative charge is in the opposite direction compared to the field. This follows from the definition, since dividing a vector by a negative number reverses its direction. It’s as though we had some objects that fell upward instead of down.

**self-check A**

Find an equation for the magnitude of the field of a single point charge $Q$.  

▷ Answer, p. 1039
Example 3.

Superposition of electric fields

Charges \( q \) and \( -q \) are at a distance \( b \) from each other, as shown in the figure. What is the electric field at the point \( P \), which lies at a third corner of the square?

The field at \( P \) is the vector sum of the fields that would have been created by the two charges independently. Let positive \( x \) be to the right and let positive \( y \) be up.

Negative charges have fields that point at them, so the charge \( -q \) makes a field that points to the right, i.e., has a positive \( x \) component. Using the answer to the self-check, we have

\[
E_{-q,x} = \frac{kq}{b^2} \\
E_{-q,y} = 0.
\]

Note that if we had blindly ignored the absolute value signs and plugged in \(-q\) to the equation, we would have incorrectly concluded that the field went to the left.

By the Pythagorean theorem, the positive charge is at a distance \( \sqrt{2}b \) from \( P \), so the magnitude of its contribution to the field is \( E = \frac{kq}{2b^2} \). Positive charges have fields that point away from them, so the field vector is at an angle of 135° counterclockwise from the \( x \) axis.

\[
E_{q,x} = \frac{kq}{2b^2} \cos 135° \\
= -\frac{kq}{2^{3/2}b^2} \\
E_{q,y} = \frac{kq}{2b^2} \sin 135° \\
= \frac{kq}{2^{3/2}b^2}
\]

The total field is

\[
E_x = \left(1 - 2^{-3/2}\right) \frac{kq}{b^2} \\
E_y = \frac{kq}{2^{3/2}b^2}.
\]
A line of charge example 4

In a complete circuit, there is typically no net charge on any of the wires. However, there are some devices in which a circuit is intentionally left open in order to produce a nonzero net charge on a wire. One example is a type of radiation detector called a Geiger-Müller tube, figure j. A high voltage is applied between the outside of the cylinder and the wire that runs along the central axis. A net positive charge builds up on the wire and a negative one on the cylinder’s wall. Electric fields originate from the wire, spread outward from the axis, and terminate on the wall. The cylinder is filled with a low-pressure inert gas. An incoming particle of radioactivity strikes an atom of the gas, ionizing it, i.e., splitting it into positively and negatively charged parts, known as ions. These ions then accelerate in opposite directions, since the force exerted by an electric field on a charged particle flips directions when the charge is reversed. The ions accelerate up to speeds at which they are capable of ionizing other atoms when they collide with them. The result is an electrical avalanche that causes a disturbance on the voltmeter.

Motivated by this example, we would like to find how the field of a long, uniformly charged wire varies with distance. In figure k/1, the point P experiences a field that is the vector sum of contributions such as the one coming from the segment q. The field $E_q$ arising from this segment has to be added to similar contributions from all other segments of the wire. By symmetry, the total field will end up pointing at a right angle to the wire. We now consider point $P'$, figure k/2, at twice the distance from the wire. If we reproduce all the angles from k/1, then the new triangle is simply a copy of the old one that has been scaled up by a factor of two. The left side’s length has doubled, so $q' = 2q$, and this would tend to make $E_{q'}$ twice as big. But all the distances have also been doubled, and the $1/r^2$ in Coulomb’s law therefore contributes an additional factor of 1/4. Combining these two factors, we find $E_{q'} = E_q/2$. The total field is the sum of contributions such as $E_{q'}$, so if all of these have been weakened by a factor of two, the same must apply to the total as well. There was nothing special about the number 2, so we conclude that in general the electric field of a line of charge is proportional to $1/r$.

Applying this to the Geiger-Müller tube, we can see the reason why the device is built with a wire. When $r$ is small, $1/r$ is big, and the field is very strong. Therefore the device can be sensitive enough to trigger an avalanche in the gas when only a single atom has been ionized.

We have only shown that the field is proportional to $1/r$, but we haven’t filled in the other factors in the equation. This is done in example 15 on p. 655.
Dipoles

The simplest set of sources that can occur with electricity but not with gravity is the dipole, consisting of a positive charge and a negative charge with equal magnitudes. More generally, an electric dipole can be any object with an imbalance of positive charge on one side and negative on the other. A water molecule, \( n \), is a dipole because the electrons tend to shift away from the hydrogen atoms and onto the oxygen atom.

Your microwave oven acts on water molecules with electric fields. Let us imagine what happens if we start with a uniform electric field, \( l/1 \), made by some external charges, and then insert a dipole, \( l/2 \), consisting of two charges connected by a rigid rod. The dipole disturbs the field pattern, but more important for our present purposes is that it experiences a torque. In this example, the positive charge feels an upward force, but the negative charge is pulled down. The result is that the dipole wants to align itself with the field, \( l/3 \). The microwave oven heats food with electrical (and magnetic) waves. The alternation of the torque causes the molecules to wiggle and increase the amount of random motion. The slightly vague definition of a dipole given above can be improved by saying that a dipole is any object that experiences a torque in an electric field.

What determines the torque on a dipole placed in an externally created field? Torque depends on the force, the distance from the axis at which the force is applied, and the angle between the force and the line from the axis to the point of application. Let a dipole consisting of charges \( +q \) and \( -q \) separated by a distance \( \ell \) be placed in an external field of magnitude \( |E| \), at an angle \( \theta \) with respect to the field. The total torque on the dipole is

\[
\tau = \ell q |E| \sin \theta.
\]

(Note that even though the two forces are in opposite directions, the torques do not cancel, because they are both trying to twist the
dipole in the same direction.) The quantity \( \ell q \) is called the dipole moment, notated \( D \).

**Dipole moment of a molecule of NaCl gas**  
In a molecule of NaCl gas, the center-to-center distance between the two atoms is about 0.24 nm. Assuming that the chlorine completely steals one of the sodium’s electrons, compute the magnitude of this molecule’s dipole moment.

The total charge is zero, so it doesn’t matter where we choose the origin of our coordinate system. For convenience, let’s choose it to be at one of the atoms, so that the charge on that atom doesn’t contribute to the dipole moment. The magnitude of the dipole moment is then

\[
D = \ell q = (2.4 \times 10^{-10} \text{ m})(e) = (2.4 \times 10^{-10} \text{ m})(1.6 \times 10^{-19} \text{ C}) \\
\approx 4 \times 10^{-29} \text{ C} \cdot \text{m}.
\]

The experimentally measured value is \(3.0 \times 10^{-29} \text{ C} \cdot \text{m} \), which shows that the electron is not completely “stolen.”

More complex dipoles can also be assigned a dipole moment — they are defined as having the same dipole moment as the two-charge dipole that would experience the same torque.

**Molecules with zero and nonzero dipole moments**  
It can be useful to know whether or not a molecule is polar, i.e., has a nonzero dipole moment. A polar molecule such as water is readily heated in a microwave oven, while a nonpolar one is not. Polar molecules are attracted to one another, so polar substances dissolve in other polar substances, but not in nonpolar substances, i.e., “like dissolves like.”

In a symmetric molecule such as carbon disulfide, figure o/1, the dipole moment vanishes. For if we rotate the molecule by 180 degrees about any one of the three coordinate axes defined in the caption of the figure, the molecule is unchanged, which means that its dipole moment is unchanged. This means that \( \text{CS}_2 \) cannot be equivalent to any simple, two-charge dipole, because a simple dipole can only stay the same under a 180 degree rotation if the rotation is about the line connecting the two charges.
Similar symmetry arguments show that sulfur hexafluoride, $\text{o/2}$, and benzene $\text{o/3}$, have vanishing dipole moments.

The formaldehyde molecule, $\text{o/4}$, does not have enough symmetry to guarantee that its dipole moment must vanish, but it does have enough to dictate it must be equivalent to a two-charge dipole lying along the left-right axis. Chloroform, $\text{o/5}$, is a front-to-back dipole for the orientation drawn in the figure.

From these considerations we can tell, for example, that carbon disulfide will be soluble in benzene, but chloroform will not.

Symmetry arguments are not enough to determine, for example, whether formaldehyde is equivalent to a dipole whose positive charge lies to the left of its negative one or to the right. This requires some knowledge of chemistry and the periodic table.

**Alternative definition of the electric field**

The behavior of a dipole in an externally created field leads us to an alternative definition of the electric field:

**alternative definition of the electric field**

The electric field vector, $\mathbf{E}$, at any location in space is defined by observing the torque exerted on a test dipole $D_t$ placed there. The direction of the field is the direction in which the field tends to align a dipole (from $-$ to $+$), and the field’s magnitude is $|\mathbf{E}| = \tau / D_t \sin \theta$.

The main reason for introducing a second definition for the same concept is that the magnetic field is most easily defined using a similar approach.

**Potential related to electric field**

Voltage (electric potential) is potential energy per unit charge, and electric field is force per unit charge. We can therefore relate potential and field if we start from the relationship between potential energy and force,

\[
\Delta PE = -Fd, \quad \text{[assuming constant force and motion parallel to the force]}
\]

and divide by charge,

\[
\frac{\Delta PE}{q} = -\frac{F}{q} d,
\]

giving

\[
\Delta V = -Ed, \quad \text{[assuming constant force and motion parallel to the force]}
\]

In other words, the difference in potential between two points equals the electric field strength multiplied by the distance between them.
The interpretation is that a strong electric field is a region of space where the potential is rapidly changing. By analogy, a steep hillside is a place on the map where the altitude is rapidly changing.

Field generated by an electric eel example 7

Suppose an electric eel is 1 m long, and generates a voltage difference of 1000 volts between its head and tail. What is the electric field in the water around it?

We are only calculating the amount of field, not its direction, so we ignore positive and negative signs. Subject to the possibly inaccurate assumption of a constant field parallel to the eel's body, we have

\[ |E| = \frac{\Delta V}{\Delta x} = 1000 \text{ V/m}. \]

The hammerhead shark example 8

One of the reasons hammerhead sharks have their heads shaped the way they do is that, like quite a few other fish, they can sense electric fields as a way of finding prey, which may for example be hidden in the sand. From the equation \( E = \frac{\Delta V}{\Delta x} \), we can see that if the shark is sensing the potential difference between two points, it will be able to detect smaller electric fields if those two points are farther apart. The shark has a network of sensory organs, called the ampullae of Lorenzini, on the skin of its head. Since the network is spread over a wider head, the \( \Delta x \) is larger. Some sharks can detect electric fields as weak as 50 picovolts per meter!

Relating the units of electric field and potential example 9

From our original definition of the electric field, we expect it to have units of newtons per coulomb, N/C. The example above, however, came out in volts per meter, V/m. Are these inconsistent? Let's reassure ourselves that this all works. In this kind of situation, the best strategy is usually to simplify the more complex units so that they involve only mks units and coulombs. Since potential is defined as electrical energy per unit charge, it has units of J/C:

\[ \frac{V}{m} = \frac{J}{C \cdot m} = \frac{J}{m} = \frac{N \cdot m}{C \cdot m}. \]

To connect joules to newtons, we recall that work equals force times distance, so \( J = N \cdot m \), so

\[ \frac{V}{m} = \frac{N \cdot m}{C \cdot m} = \frac{N}{C}. \]

As with other such difficulties with electrical units, one quickly begins to recognize frequently occurring combinations.
Discussion questions

A  In the definition of the electric field, does the test charge need to be 1 coulomb? Does it need to be positive?

B  Does a charged particle such as an electron or proton feel a force from its own electric field?

C  Is there an electric field surrounding a wall socket that has nothing plugged into it, or a battery that is just sitting on a table?

D  In a flashlight powered by a battery, which way do the electric fields point? What would the fields be like inside the wires? Inside the filament of the bulb?

E  Criticize the following statement: “An electric field can be represented by a sea of arrows showing how current is flowing.”

F  The field of a point charge, \( |E| = \frac{kQ}{r^2} \), was derived in the self-check above. How would the field pattern of a uniformly charged sphere compare with the field of a point charge?

G  The interior of a perfect electrical conductor in equilibrium must have zero electric field, since otherwise the free charges within it would be drifting in response to the field, and it would not be in equilibrium. What about the field right at the surface of a perfect conductor? Consider the possibility of a field perpendicular to the surface or parallel to it.

H  Compare the dipole moments of the molecules and molecular ions shown in the figure.

I  Small pieces of paper that have not been electrically prepared in any way can be picked up with a charged object such as a charged piece of tape. In our new terminology, we could describe the tape’s charge as inducing a dipole moment in the paper. Can a similar technique be used to induce not just a dipole moment but a charge?

J  The earth and moon are fairly uneven in size and far apart, like a baseball and a ping-pong ball held in your outstretched arms. Imagine instead a planetary system with the character of a double planet: two planets of equal size, close together. Sketch a sea of arrows diagram of their gravitational field.
22.4 Calculating energy in fields

We found on p. 634 that fields have energy, and we now know as well that fields act like vectors. Presumably there is a relationship between the strength of a field and its energy density. Flipping the direction of the field can’t change the density of energy, which is a scalar and therefore has no direction in space. We therefore expect that the energy density to be proportional to the square of the field, so that changing $E$ to $-E$ has no effect on the result. This is exactly what we’ve already learned to expect for waves: the energy depends on the square of the amplitude. The relevant equations for the gravitational and electric fields are as follows:

\[
\text{(energy stored in the gravitational field per m}^3) = -\frac{1}{8\pi G} |g|^2
\]

\[
\text{(energy stored in the electric field per m}^3) = \frac{1}{8\pi k} |E|^2
\]

A similar expression is given on p. 701 for the magnetic field.

Although funny factors of $8\pi$ and the plus and minus signs may have initially caught your eye, they are not the main point. The important idea is that the energy density is proportional to the square of the field strength in all cases. We first give a simple numerical example and work a little on the concepts, and then turn our attention to the factors out in front.

In chapter 22 when we discussed the original reason for introducing the concept of a field of force, a prime motivation was that otherwise there was no way to account for the energy transfers involved when forces were delayed by an intervening distance. We used to think of the universe’s energy as consisting of

kinetic energy  
+gravitational potential energy based on the distances between objects that interact gravitationally  
+electric potential energy based on the distances between objects that interact electrically  
+magnetic potential energy based on the distances between objects that interact magnetically,

but in nonstatic situations we must use a different method:

kinetic energy  
+gravitational potential energy stored in gravitational fields  
+electric potential energy stored in electric fields  
+magnetic potential stored in magnetic fields
Surprisingly, the new method still gives the same answers for the static cases.

**Energy stored in a capacitor example 10**

A pair of parallel metal plates, seen from the side in figure r, can be used to store electrical energy by putting positive charge on one side and negative charge on the other. Such a device is called a capacitor. (We have encountered such an arrangement previously, but there its purpose was to deflect a beam of electrons, not to store energy.)

In the old method of describing potential energy, 1, we think in terms of the mechanical work that had to be done to separate the positive and negative charges onto the two plates, working against their electrical attraction. The new description, 2, attributes the storage of energy to the newly created electric field occupying the volume between the plates. Since this is a static case, both methods give the same, correct answer.

**Potential energy of a pair of opposite charges example 11**

Imagine taking two opposite charges, s, that were initially far apart and allowing them to come together under the influence of their electrical attraction.

According to the old method, potential energy is lost because the electric force did positive work as it brought the charges together. (This makes sense because as they come together and accelerate it is their potential energy that is being lost and converted to kinetic energy.)

By the new method, we must ask how the energy stored in the electric field has changed. In the region indicated approximately by the shading in the figure, the superposing fields of the two charges undergo partial cancellation because they are in opposing directions. The energy in the shaded region is reduced by this effect. In the unshaded region, the fields reinforce, and the energy is increased.

It would be quite a project to do an actual numerical calculation of the energy gained and lost in the two regions (this is a case where the old method of finding energy gives greater ease of computation), but it is fairly easy to convince oneself that the energy is less when the charges are closer. This is because bringing the charges together shrinks the high-energy unshaded region and enlarges the low-energy shaded region.
Energy transmitted by ripples in the electric and magnetic fields

**example 12**

We’ll see in chapter 24 that phenomena like light, radio waves, and x-rays are all ripples in the electric and magnetic fields. The old method would give zero energy for a region of space containing a light wave but no charges. That would be wrong! We can only use the old method in static cases.

Now let’s give at least some justification for the other features of the expressions for energy density, $-\frac{1}{8\pi G}g^2$ and $\frac{1}{8\pi k}E^2$, besides the proportionality to the square of the field strength.

First, why the different plus and minus signs? The basic idea is that the signs have to be opposite in the gravitational and electric cases because there is an attraction between two positive masses (which are the only kind that exist), but two positive charges would repel. Since we’ve already seen examples where the positive sign in the electric energy makes sense, the gravitational energy equation must be the one with the minus sign.

It may also seem strange that the constants $G$ and $k$ are in the denominator. They tell us how strong the three different forces are, so shouldn’t they be on top? No. Consider, for instance, an alternative universe in which gravity is twice as strong as in ours. The numerical value of $G$ is doubled. Because $G$ is doubled, all the gravitational field strengths are doubled as well, which quadruples the quantity $|g|^2$. In the expression $-\frac{1}{8\pi G}|g|^2$, we have quadrupled something on top and doubled something on the bottom, which makes the energy twice as big. That makes perfect sense.

**Discussion questions**

A The figure shows a positive charge in the gap between two capacitor plates. First make a large drawing of the field pattern that would be formed by the capacitor itself, without the extra charge in the middle. Next, show how the field pattern changes when you add the particle at these two positions. Compare the energy of the electric fields in the two cases. Does this agree with what you would have expected based on your knowledge of electrical forces?

B Criticize the following statement: “A solenoid makes a charge in the space surrounding it, which dissipates when you release the energy.”

C In example 11, I argued that the fields surrounding a positive and negative charge contain less energy when the charges are closer together. Perhaps a simpler approach is to consider the two extreme possibilities: the case where the charges are infinitely far apart, and the one in which they are at zero distance from each other, i.e., right on top of each other. Carry out this reasoning for the case of (1) a positive charge and a negative charge of equal magnitude, (2) two positive charges of equal magnitude, (3) the gravitational energy of two equal masses.
22.5 Potential for nonuniform fields

The calculus-savvy reader will have no difficulty generalizing the field-potential relationship to the case of a varying field. The potential energy associated with a varying force is

\[ \Delta PE = - \int F \, dx, \quad \text{[one dimension]} \]

so for electric fields we divide by \( q \) to find

\[ \Delta V = - \int E \, dx, \quad \text{[one dimension]} \]

Applying the fundamental theorem of calculus yields

\[ E = - \frac{dV}{dx}. \quad \text{[one dimension]} \]

Potential associated with a point charge example 13

What is the potential associated with a point charge?

As derived previously in self-check A on page 639, the field is

\[ |E| = \frac{kQ}{r^2} \]

The difference in potential between two points on the same radius line is

\[ \Delta V = \int dV = - \int E_x \, dx \]

In the general discussion above, \( x \) was just a generic name for distance traveled along the line from one point to the other, so in this case \( x \) really means \( r \).

\[ \Delta V = - \int_{r_1}^{r_2} E_r \, dr \]

\[ = - \int_{r_1}^{r_2} \frac{kQ}{r^2} \, dr \]

\[ = \left[ \frac{kQ}{r} \right]_{r_1}^{r_2} \]

\[ = kQ \left( \frac{1}{r_2} - \frac{1}{r_1} \right). \]

The standard convention is to use \( r_1 = \infty \) as a reference point, so that the potential at any distance \( r \) from the charge is

\[ V = \frac{kQ}{r}. \]