band, a mobile hole, and turns the aluminum atom into a fixed negative ion. Thanks to the holes thus created – at room temperature nearly equal in number to the aluminum atoms added – the crystal becomes a much better conductor. There are also a few electrons in the conduction band, but the overwhelming majority of the mobile charge carriers are positive, and we call this material a *p-type semiconductor* (Fig. 4.11(b)).

Once the number of mobile charge carriers has been established, whether electrons or holes or both, the conductivity depends on their mobility, which is limited, as in metallic conduction, by scattering within the crystal. A single homogeneous semiconductor obeys Ohm's law. The spectacularly nonohmic behavior of semiconductor devices – as in a rectifier or a transistor – is achieved by combining n-type material with p-type material in various arrangements.

Example (Mean free time in silicon) In Fig. 4.10, a conductivity of 30 (ohm-m)$^{-1}$ results from the presence of 10^{21} electrons per m^3 in the conduction band, along with the same number of holes. Assume that $\tau_+ = \tau_-$ and $M_+ = M_- = m_e$, the electron mass. What must be the value of the mean free time τ? The rms speed of an electron at 500 K is $1.5 \cdot 10^5$ m/s. Compare the mean free path with the distance between neighboring silicon atoms, which is $2.35 \cdot 10^{-10}$ m.

Solution Since we have two types of charge carriers, the electrons and the holes, Eq. (4.23) gives

$$\tau = \frac{m\sigma}{2Ne^2} = \frac{(9.1 \cdot 10^{-31}\ \text{kg})\big(30\ (\text{ohm-m})^{-1}\big)}{2(10^{21}\ \text{m}^{-3})(1.6 \cdot 10^{-19}\ \text{C})^2} \approx 5.3 \cdot 10^{-13}\ \text{s}. \quad (4.28)$$

The distance traveled during this time is $v\tau = (1.5 \cdot 10^5\ \text{m/s})(5.3 \cdot 10^{-13}\ \text{s}) \approx 8 \cdot 10^{-8}$ m, which is more than 300 times the distance between neighboring silicon atoms.

4.7 Circuits and circuit elements

Electrical devices usually have well-defined terminals to which wires can be connected. Charge can flow into or out of the device over these paths. In particular, if two terminals, and only two, are connected by wires to something outside, and if the current flow is steady with constant potentials everywhere, then obviously the current must be equal and opposite at the two terminals.[13] In that case we can speak of *the* current I that flows through the device, and of *the* voltage V "between the terminals" or "across the terminals," which means their difference in electric

[13] It is perfectly possible to have 4 A flowing into one terminal of a two-terminal object with 3 A flowing out at the other terminal. But then the object is accumulating positive charge at the rate of 1 coulomb/second. Its potential must be changing very rapidly – and that can't go on for long. Hence this cannot be a *steady,* or time-independent, current.

potential. The ratio V/I for some given I is a certain number of resistance units (ohms, if V is in volts and I in amps). If Ohm's law is obeyed in all parts of the object through which current flows, that number will be a constant, independent of the current. This one number completely describes the electrical behavior of the object, for steady current flow (DC) between the given terminals. With these rather obvious remarks we introduce a simple idea, the notion of a *circuit element.*

Look at the five boxes in Fig. 4.12. Each has two terminals, and inside each box there is some stuff, different in every box. If any one of these boxes is made part of an electrical circuit by connecting wires to the terminals, the ratio of the potential difference between the terminals to the current flowing in the wire that we have connected to the terminal will be found to be 65 ohms. We say the resistance between the terminals, in each box, is 65 ohms. This statement would surely not be true for all conceivable values of the current or potential difference. As the potential difference or *voltage* between the terminals is raised, various things might happen, earlier in some boxes than in others, to change the *voltage/current* ratio. You might be able to guess which boxes would give trouble first. Still, there is *some* limit below which they all behave linearly; within that range, for *steady* currents, the boxes are alike. They are alike in this sense: if any circuit contains one of these boxes, which box it is makes no difference in the behavior of that circuit. The box is equivalent to a 65 ohm resistor.[14] We represent it by the symbol –∧∧∧– and in the description of the circuit of which the box is one component, we replace the box with this abstraction. An electrical circuit or network is then a collection of such circuit elements joined to one another by paths of negligible resistance.

Taking a network consisting of many elements connected together and selecting two points as terminals, we can regard the whole thing as equivalent, as far as these two terminals are concerned, to a single resistor. We say that the physical network of objects in Fig. 4.13(a) is represented by the diagram of Fig. 4.13(b), and for the terminals A_1A_2 the equivalent circuit is Fig. 4.13(c). The equivalent circuit for the terminals at B_1B_2 is given in Fig. 4.13(d). If you put this assembly in a box with only that pair of terminals accessible, it will be indistinguishable from a resistor of 57.6 ohm resistance.

There is one very important rule – only *direct-current* measurements are allowed! All that we have said depends on the current and electric fields being constant in time; if they are not, the behavior of a circuit

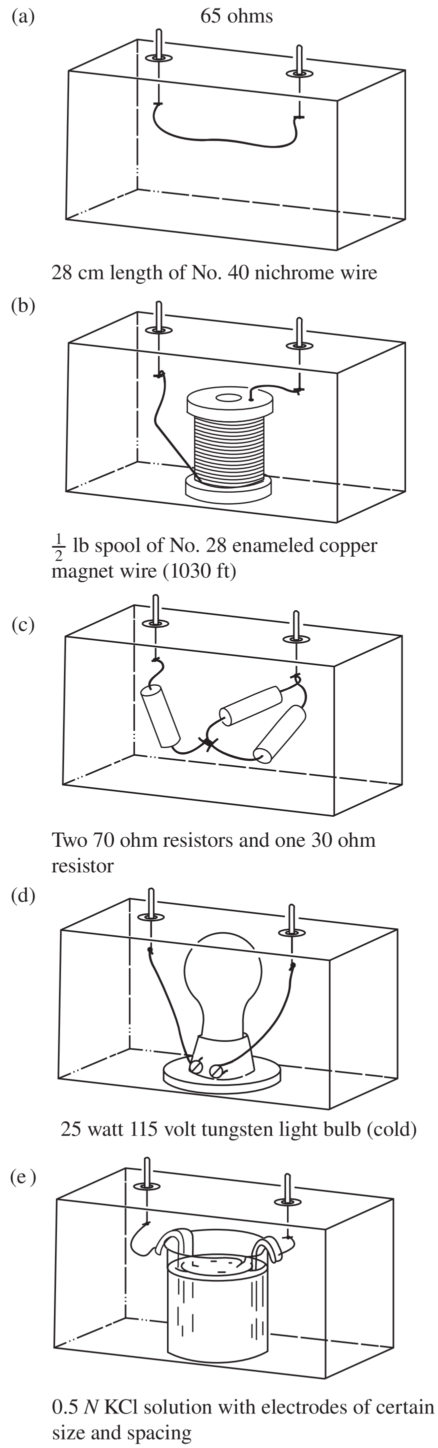

Figure 4.12.
Various devices that are equivalent, for direct current, to a 65 ohm resistor.

[14] We use the term *resistor* for the actual object designed especially for that function. Thus a "200 ohm, 10 watt, wire-wound resistor" is a device consisting of a coil of wire on some insulating base, with terminals, intended to be used in such a way that the average power dissipated in it is not more than 10 watts.

element may not depend on its resistance alone. The concept of equivalent circuits can be extended from these DC networks to systems in which current and voltage vary with time. Indeed, that is where it is most valuable. We are not quite ready to explore that domain.

Little time will be spent here on methods for calculating the equivalent resistance of a network of circuit elements. The cases of series and parallel groups are easy. A combination like that in Fig. 4.14 is two resistors, of value R_1 and R_2, in series. The equivalent resistance is

$$\boxed{R = R_1 + R_2} \tag{4.29}$$

A combination like that in Fig. 4.15 is two resistors in parallel. By an argument that you should be able to give (see Problem 4.3), the equivalent resistance R is found to be

$$\boxed{\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}} \quad \text{or} \quad R = \frac{R_1 R_2}{R_1 + R_2}. \tag{4.30}$$

Example (Reducing a network) Let's use the addition rules in Eqs. (4.29) and (4.30) to reduce the network shown in Fig. 4.16 to an equivalent single resistor. As complicated as this network looks, it can be reduced, step by step, via series or parallel combinations. We assume that every resistor in the circuit has the value 100 ohms.

Using the above rules, we can reduce the network as follows (you should verify all of the following statements). A parallel combination of two 100 ohm resistors is equivalent to 50 ohms. So in the first figure, the top two circled sections are each equivalent to 150 ohms, and the bottom one is equivalent to 50 ohms. In the second figure, the top and bottom circled sections are then equivalent to 160 ohms and 150 ohms. In the third figure, the circled section is then equivalent to 77.4 ohms. The whole circuit is therefore equivalent to $100 + 77.4 + 150 = 327.4$ ohms.

Although Eqs. (4.29) and (4.30) are sufficient to handle the complicated circuit in Fig. 4.16, the simple network of Fig. 4.17 *cannot* be so reduced, so a more general method is required (see Exercise 4.44). Any conceivable network of resistors in which a constant current is flowing has to satisfy these conditions (the first is Ohm's law, the second and third are known as Kirchhoff's rules):

(1) The current through each element must equal the voltage across that element divided by the resistance of the element.

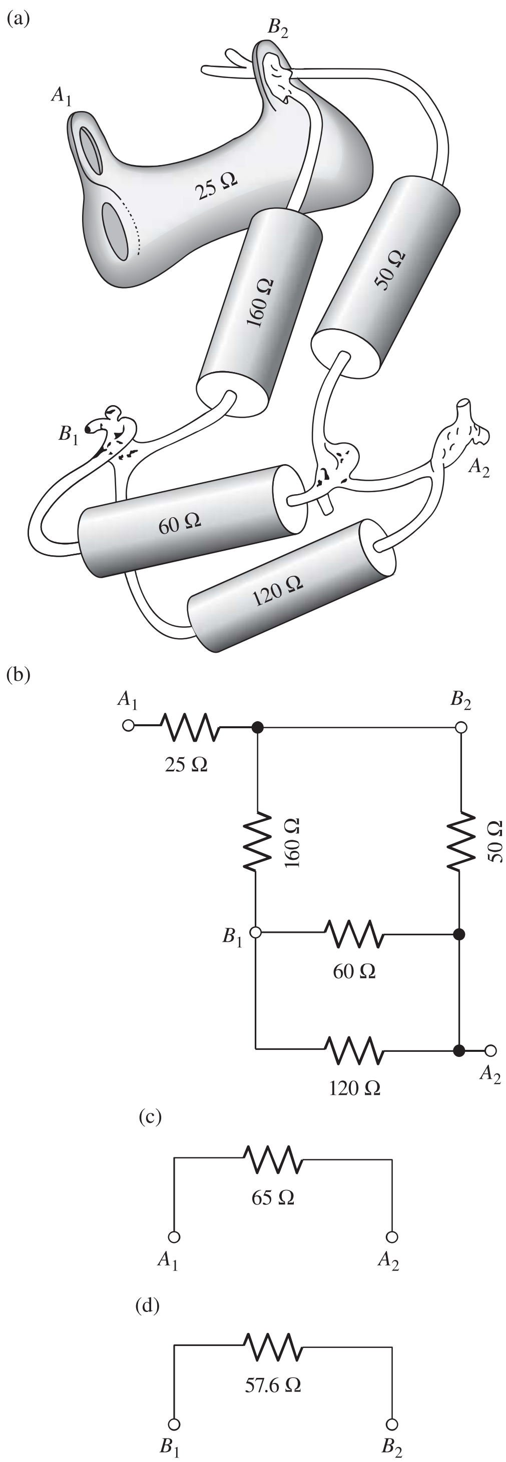

Figure 4.13.
Some resistors connected together (a); the circuit diagram (b); and the equivalent resistance between certain pairs of terminals (c) and (d).

(2) At a *node* of the network, a point where three or more connecting wires meet, the algebraic sum of the currents into the node must be zero. (This is our old charge-conservation condition, Eq. (4.8), in circuit language.)

(3) The sum of the potential differences taken in order around a *loop* of the network, a path beginning and ending at the same node, is zero. (This is network language for the general property of the static electric field: $\int \mathbf{E} \cdot d\mathbf{s} = 0$ for any closed path.)

The algebraic statement of these conditions for any network will provide exactly the number of independent linear equations needed to ensure that there is one and only one solution for the equivalent resistance between two selected nodes. We assert this without proving it. It is interesting to note that the structure of a DC network problem depends only on the *topology* of the network, that is, on those features of the diagram of connections that are independent of any distortion of the lines of the diagram. We will give an example of the use of the above three rules in Section 4.10, after we have introduced the concept of electromotive force.

A DC network of resistances is a *linear* system – the voltages and currents are governed by a set of linear equations, the statements of the conditions (1), (2), and (3). Therefore the superposition of different possible states of the network is also a possible state. Figure 4.18 shows a section of a network with certain currents, $I_1, I_2, \dots,$ flowing in the wires and certain potentials, $V_1, V_2, \dots,$ at the nodes. If some other set of currents and potentials, say $I'_1, \dots, V'_1, \dots,$ is another possible state of affairs in this section of network, then so is the set $(I_1 + I'_1), \dots, (V_1 + V'_1), \dots.$ These currents and voltages corresponding to the superposition will also satisfy the conditions (1), (2), and (3). Some general theorems about networks, interesting and useful to the electrical engineer, are based on this. One such theorem is *Thévenin's theorem*, discussed in Section 4.10 and proved in Problem 4.13.

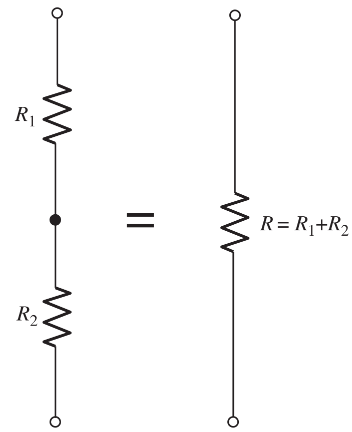

Figure 4.14.
Resistances in series.

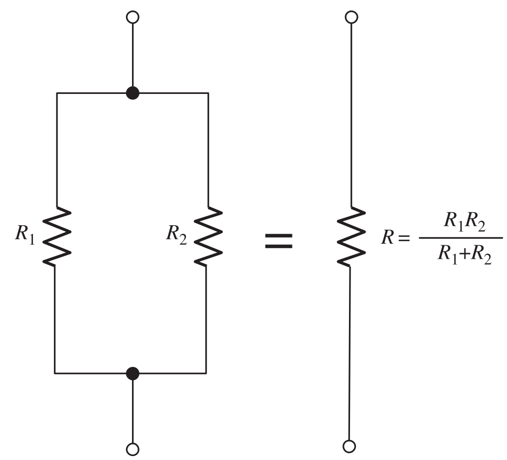

Figure 4.15.
Resistances in parallel.

4.8 Energy dissipation in current flow

The flow of current in a resistor involves the dissipation of energy. If it takes a force $\mathbf{F}$ to push a charge carrier along with average velocity $\mathbf{v}$, any agency that accomplishes this must do work at the rate $\mathbf{F} \cdot \mathbf{v}$. If an electric field $\mathbf{E}$ is driving the ion of charge q, then $\mathbf{F} = q\mathbf{E}$, and the rate at which work is done is $q\mathbf{E} \cdot \mathbf{v}$. The energy thus expended shows up eventually as heat. In our model of ionic conduction, the way this comes about is quite clear. The ion acquires some extra kinetic energy, as well as momentum, between collisions. A collision, or at most a few collisions, redirects its momentum at random but does not necessarily restore the kinetic energy to normal. For that to happen the ion has to transfer kinetic energy to the obstacle that deflects it. Suppose the charge carrier has a considerably smaller mass than the neutral atom it collides with. The average transfer

4.9 Electromotive force and the voltaic cell

The origin of the electromotive force in a direct-current circuit is some mechanism that transports charge carriers in a direction *opposite* to that in which the electric field is trying to move them. A Van de Graaff electrostatic generator (Fig. 4.19) is an example on a large scale. With everything running steadily, we find current in the external resistance flowing in the direction of the electric field **E**, and energy being dissipated there (appearing as heat) at the rate IV_0, or I^2R. Inside the column of the machine, too, there is a downward-directed electric field. Here charge carriers can be moved against the field if they are stuck to a nonconducting belt. They are stuck so tightly that they can't slide backward along the belt in the generally downward electric field. (They can still be removed from the belt by a much stronger field localized at the brush in the terminal. We need not consider here the means for putting charge on and off the belt near the pulleys.) The energy needed to pull the belt is supplied from elsewhere – usually by an electric motor connected to a power line, but it could be a gasoline engine, or even a person turning a crank. This Van de Graaff generator is in effect a battery with an electromotive force, under these conditions, of V_0 volts.

In ordinary batteries it is chemical energy that makes the charge carriers move through a region where the electric field opposes their motion. That is, a *positive* charge carrier may move to a place of *higher* electric potential if by so doing it can engage in a chemical reaction that will yield more energy than it costs to climb the electrical hill.

To see how this works, let us examine one particular voltaic cell. *Voltaic cell* is the generic name for a chemical source of electromotive force. In the experiments of Galvani around 1790 the famous twitching frogs' legs had signaled the chemical production of electric current. It was Volta who proved that the source was not "animal electricity," as Galvani maintained, but the contact of dissimilar metals in the circuit. Volta went on to construct the first battery, a stack of elementary cells, each of which consisted of a zinc disk and a silver disk separated by cardboard moistened with brine. The battery that powers your flashlight comes in a tidier package, but the principle of operation is the same. Several kinds of voltaic cells are in use, differing in their chemistry but having common features: two electrodes of different material immersed in an ionized fluid, or electrolyte.

As an example, we'll describe the lead–sulfuric acid cell which is the basic element of the automobile battery. This cell has the important property that its operation is readily reversible. With a *storage battery* made of such cells, which can be charged and discharged repeatedly, energy can be stored and recovered electrically.

A fully charged lead–sulfuric acid cell has positive plates that hold lead dioxide, PbO_2, as a porous powder, and negative plates that hold pure lead of a spongy texture. The mechanical framework, or grid, is made of a lead alloy. All the positive plates are connected together and

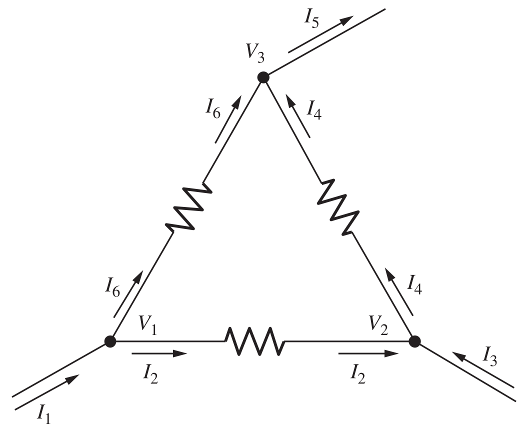

Figure 4.18.
Currents and potentials at the nodes of a network.

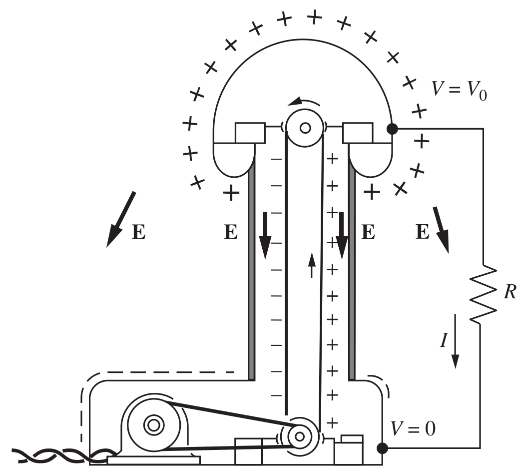

Figure 4.19.
In the Van de Graaff generator, charge carriers are mechanically transported in a direction opposite to that in which the electric field would move them.

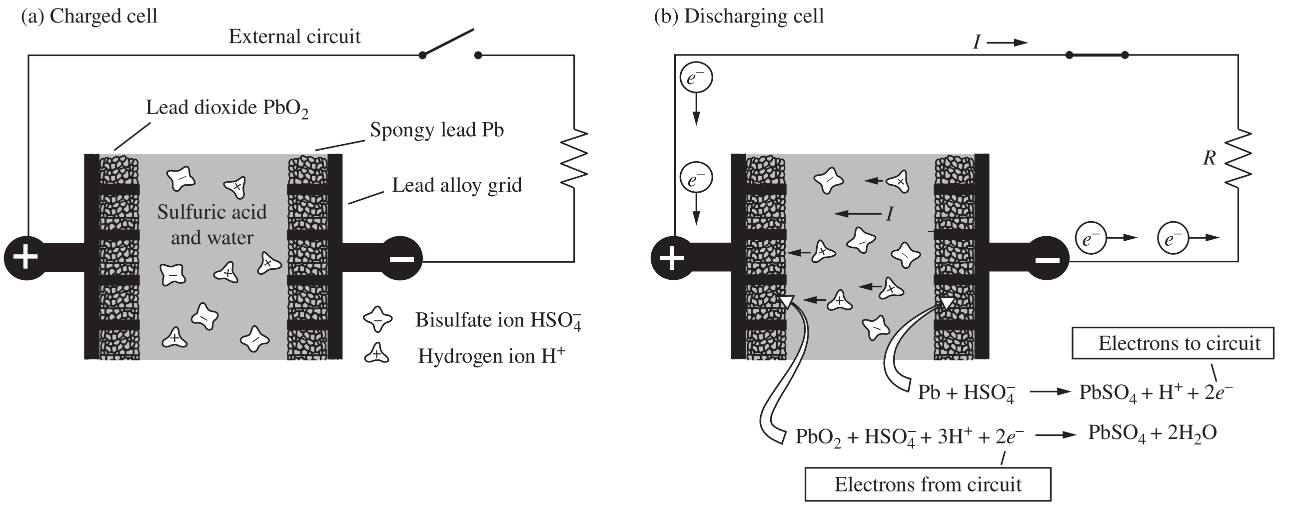

Figure 4.20.
A schematic diagram, not to scale, showing how the lead–sulfuric acid cell works. The electrolyte, sulfuric acid solution, permeates the lead oxide granules in the positive plate and the spongy lead in the negative plate. The potential difference between the positive and negative terminals is 2.1 V. With the external circuit closed, chemical reactions proceed at the solid–liquid interfaces in both plates, resulting in the depletion of sulfuric acid in the electrolyte and the transfer of electrons through the external circuit from negative terminal to positive terminal, which constitutes the current I. To recharge the cell, replace the load R by a source with electromotive force greater than 2.1 V, thus forcing current to flow through the cell in the opposite direction and reversing both reactions.

to the positive terminal of the cell. The negative plates, likewise connected, are interleaved with the positive plates, with a small separation. The schematic diagram in Fig. 4.20 shows only a small portion of a positive and a negative plate. The sulfuric acid electrolyte fills the cell, including the interstices of the active material, the porosity of which provides a large surface area for chemical reaction.

The cell will remain indefinitely in this condition if there is no external circuit connecting its terminals. The potential difference between its terminals will be close to 2.1 volts. This open-circuit potential difference is established "automatically" by the chemical interaction of the constituents. This is the *electromotive force* of the cell, for which the symbol $\mathcal{E}$ will be used. Its value depends on the concentration of sulfuric acid in the electrolyte, but not at all on the size, number, or separation of the plates.

Now connect the cell's terminals through an external circuit with resistance R. If R is not too small, the potential difference V between the cell terminals will drop only a little below its open-circuit value $\mathcal{E}$, and a current $I = V/R$ will flow around the circuit (Fig. 4.20(b)). Electrons flow *into* the positive terminal; other electrons flow *out* of the negative terminal. At each electrode chemical reactions are proceeding, the overall effect of which is to convert lead, lead dioxide, and sulfuric acid into lead sulfate and water. For every molecule of lead sulfate thus made, one charge e is passed around the circuit and an amount of energy $e\mathcal{E}$ is released. Of this energy the amount eV appears as heat in the external resistance R. The difference between $\mathcal{E}$ and V is caused by the resistance of the electrolyte itself, through which the current I must flow inside the cell. If we represent this internal resistance by R_i, the system can be quite well described by the equivalent circuit in Fig. 4.21.

As discharge goes on and the electrolyte becomes more diluted with water, the electromotive force $\mathcal{E}$ decreases somewhat. Normally, the cell is considered discharged when $\mathcal{E}$ has fallen below 1.75 volts. To recharge the cell, current must be forced around the circuit in the opposite direction by connecting a voltage source greater than $\mathcal{E}$ across the cell's terminals. The chemical reactions then run backward until all the lead sulfate is turned back into lead dioxide and lead. The investment of energy in charging the cell is somewhat more than the cell will yield on discharge, for the internal resistance R_i causes a power loss I^2R_i whichever way the current is flowing.

Note in Fig. 4.20(b) that the current I in the electrolyte is produced by a net drift of positive ions toward the positive plate. Evidently the electric field in the electrolyte points toward, not away from, the positive plate. Nevertheless, the line integral of $\mathbf{E}$ around the whole circuit is zero, as it must be for any electrostatic field. The explanation is this: there are two very steep jumps in potential at the interface of the positive plate and the electrolyte and at the interface of the negative plate and the electrolyte. That is where the ions are moved *against* a strong electric field by forces arising in the chemical reactions. It is this region that corresponds to the belt in a Van de Graaff generator.

Every kind of voltaic cell has its characteristic electromotive force, falling generally in the range of 1 to 3 volts. The energy involved, per molecule, in any chemical reaction is essentially the gain or loss in the transfer of an outer electron from one atom to a different atom. That is never more than a few electron-volts. We can be pretty sure that no one is going to invent a voltaic cell with a 12 volt electromotive force. The 12 volt automobile battery consists of six separate lead–sulfuric acid cells connected in series. For more discussion of how batteries work, including a helpful analogy, see Roberts (1983).

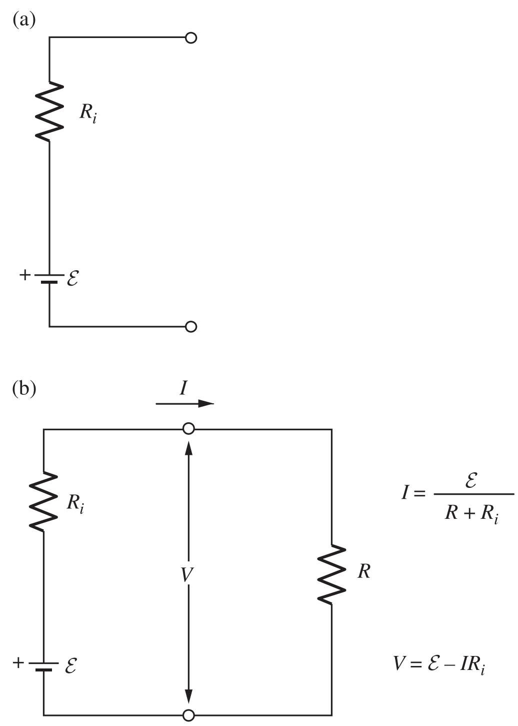

Figure 4.21.
(a) The equivalent circuit for a voltaic cell is simply a resistance R_i in series with an electromotive force $\mathcal{E}$ of fixed value.
(b) Calculation of the current in a circuit containing a voltaic cell.

Example (Lead–acid battery) A 12 V lead–acid storage battery with a 20 ampere-hour capacity rating has a mass of 10 kg.

(a) How many kilograms of lead sulfate are formed when this battery is discharged? (The molecular weight of $PbSO_4$ is 303.)
(b) How many kilograms of batteries of this type would be required to store the energy derived from 1 kg of gasoline by an engine of 20 percent efficiency? (The heat of combustion of gasoline is $4.5 \cdot 10^7$ J/kg.)

Solution

(a) The total charge transferred in 20 ampere-hours is $(20\,\text{C/s})(3600\,\text{s}) = 72{,}000\,\text{C}$. From Fig. 4.20(b), the creation of two electrons is associated with the creation of one molecule of $PbSO_4$. But also the absorption of two electrons is associated with the creation of another molecule of $PbSO_4$. So the travel of two electrons around the circuit is associated with the creation of two molecules of $PbSO_4$. The ratio is thus 1 to 1. The charge transferred per mole of $PbSO_4$ is therefore $(6 \cdot 10^{23})(1.6 \cdot 10^{-19}\,\text{C}) = 96{,}000\,\text{C}$. The above charge of 72,000 C therefore corresponds to 3/4 of a mole. Since each mole has a mass of 0.303 kg, the desired mass is about 0.23 kg.

(b) At 12 V, the energy output associated with a charge of 72,000 C is (12 J/C)(72,000 C) = 864,000 J. Also, 1 kg of gasoline burned at 20 percent efficiency yields an energy of $(0.2)(1\,\text{kg})(4.5\cdot 10^7\,\text{J/kg}) = 9\cdot 10^6\,\text{J}$. This is equivalent to $(9\cdot 10^6\,\text{J})/(8.64\cdot 10^5\,\text{J}) = 10.4$ batteries. Since each battery has a mass of 10 kg, this corresponds to 104 kg of batteries.

4.10 Networks with voltage sources

4.10.1 Applying Kirchhoff's rules

A network of resistors could contain more than one electromotive force, or voltage source. Consider the following example.

Example The circuit in Fig. 4.22 contains two batteries with electromotive force $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. In each of the conventional battery symbols shown, the longer line indicates the positive terminal. Assume that R_1 includes the internal resistance of one battery, R_2 that of the other. Supposing the resistances given, what are the currents in this network?

Solution Having assigned directions arbitrarily to the currents I_1, I_2, and I_3 in the branches, we can impose the requirements stated in Section 4.7. We have one node and two loops,[15] so we obtain three independent Kirchhoff equations:

$$\begin{aligned} I_1 - I_2 - I_3 &= 0,\\ \mathcal{E}_1 - R_1 I_1 - R_3 I_3 &= 0,\\ \mathcal{E}_2 + R_3 I_3 - R_2 I_2 &= 0. \end{aligned} \tag{4.32}$$

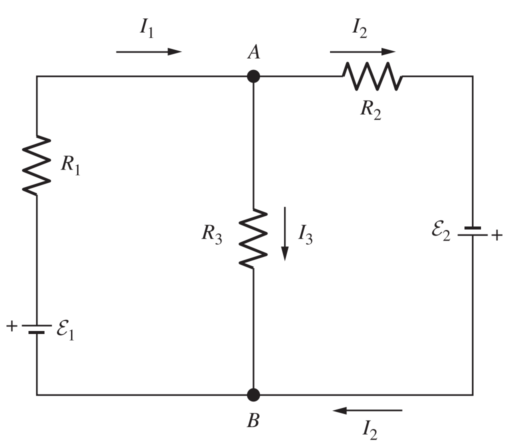

Figure 4.22.
A network with two voltage sources.

To check the signs, note that in writing the two loop equations, we have gone around each loop in the direction current would flow from the battery in that loop. The three equations can be solved for I_1, I_2, and I_3. This is slightly messy by hand, but trivial if we use a computer; the result is

$$\begin{aligned} I_1 &= \frac{\mathcal{E}_1 R_2 + \mathcal{E}_1 R_3 + \mathcal{E}_2 R_3}{R_1 R_2 + R_2 R_3 + R_1 R_3},\\ I_2 &= \frac{\mathcal{E}_2 R_1 + \mathcal{E}_2 R_3 + \mathcal{E}_1 R_3}{R_1 R_2 + R_2 R_3 + R_1 R_3},\\ I_3 &= \frac{\mathcal{E}_1 R_2 - \mathcal{E}_2 R_1}{R_1 R_2 + R_2 R_3 + R_1 R_3}. \end{aligned} \tag{4.33}$$

If in a particular case the value of I_3 turns out to be negative, it simply means that the current in that branch flows opposite to the direction we had assigned to positive current.

[15] There are actually two nodes, of course, but they give the same information. And there is technically a third loop around the whole network, but the resulting equation is the sum of the two other loop equations.

Alternatively, we can use the "loop" currents shown in Fig. 4.23. The advantages of this method are that (1) the "node" condition in Section 4.7 is automatically satisfied, because whatever current goes into a node also comes out, by construction; and (2) there are only two unknowns to solve for instead of three (although to be fair, the first of the equations in Eq. (4.32) is trivial). The disadvantage is that if we want to find the current in the middle branch (I_3 above), we need to take the difference of the loop currents I_1 and I_2, because with the sign conventions chosen, these currents pass in opposite directions through R_3. But this is not much of a burden. The two loop equations are now

$$\begin{aligned} \mathcal{E}_1 - R_1 I_1 - R_3(I_1 - I_2) &= 0, \\ \mathcal{E}_2 - R_3(I_2 - I_1) - R_2 I_2 &= 0. \end{aligned} \tag{4.34}$$

Of course, these two equations are just the second two equations in Eq. (4.32), with $I_3 = I_1 - I_2$ substituted in from the first equation. So we obtain the same values of I_1 and I_2 (and hence I_3).

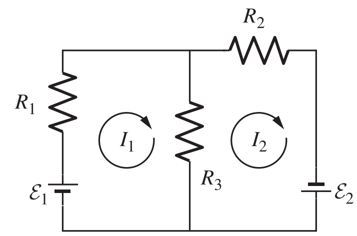

Figure 4.23.
Loop currents for use in Kirchhoff's rules. Loop currents automatically satisfy the node condition.

The calculational difference between the two methods in the above example was inconsequential. But in larger networks the second method is often more tractable, because it involves simply writing down a loop equation for every loop you see on the page. This tells you right away how many unknowns (the loop currents) there are. In either case, all of the physics is contained in the equations representing the rules given in Section 4.7. The hardest thing about these equations is making sure all the signs are correct. The actual process of solving them is easy if you use a computer. A larger network is technically no more difficult to solve than a smaller one. The only difference is that the larger network takes more time, because it takes longer to write down the equations (which are all of the same general sort) and then type them into the computer.

4.10.2 Thévenin's theorem

Suppose that a network such as the one in Fig. 4.22 forms part of some larger system, to which it is connected at two of its nodes. For example, let us connect wires to the two nodes A and B and enclose the rest in a "black box" with these two wires as the only external terminals, as in Fig. 4.24(a). A general theorem called *Thévenin's theorem* assures us that this two-terminal box is completely equivalent, in its behavior in any other circuit to which it may be connected, to a *single* voltage source $\mathcal{E}_{\text{eq}}$ ("eq" for equivalent) with an internal resistance R_{eq}. This holds for any network of voltage sources and resistors, no matter how complicated. It is not immediately obvious that such an $\mathcal{E}_{\text{eq}}$ and R_{eq} should exist (see Problem 4.13 for a proof), but assuming they do exist, their values can be determined by either experimental measurements or theoretical calculations, in the following ways.

If we *don't* know what is in the box, we can determine $\mathcal{E}_{\text{eq}}$ and R_{eq} experimentally by two measurements.

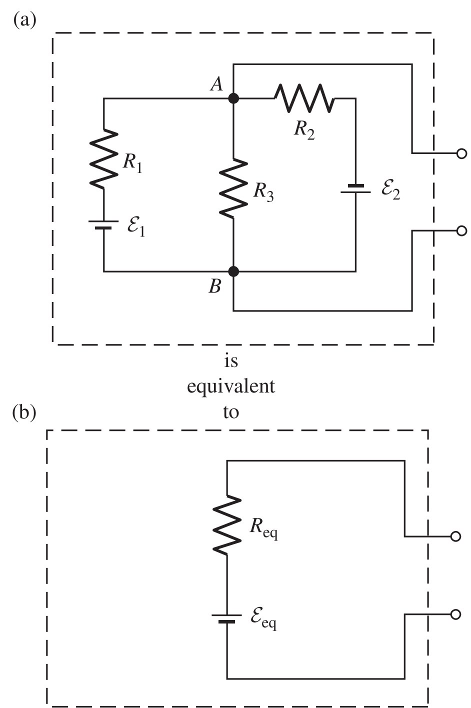

Figure 4.24.
Make R_{eq} equal to the resistance that would be measured between the terminals in (a) if all electromotive forces were zero. Make $\mathcal{E}_{\text{eq}}$ equal to the voltage observed between the terminals in (a) with the external circuit open. Then the circuit in (b) is *equivalent* to the circuit in (a). You can't tell the difference by any direct-current measurement at those terminals.

- Measure the *open-circuit voltage* between the terminals by connecting them via a voltmeter that draws negligible current. (The “infinite” resistance of the voltmeter means that the terminals are effectively unconnected; hence the name “open circuit.”) This voltage equals $\mathcal{E}_{eq}$. This is clear from Fig. 4.24(b); if essentially zero current flows through this simple circuit, then there is zero voltage drop across the resistor R_{eq}. So the measured voltage equals all of the $\mathcal{E}_{eq}$.
- Measure the *short-circuit current* I_{sc} between the terminals by connecting them via an ammeter with negligible resistance. (The “zero” resistance of the ammeter means that the terminals are effectively connected by a short circuit.) Ohm’s law for the short-circuited circuit in Fig. 4.24(b) then yields simply $\mathcal{E}_{eq} = I_{sc}R_{eq}$. The equivalent resistance is therefore given by

$$R_{eq} = \frac{\mathcal{E}_{eq}}{I_{sc}}. \qquad (4.35)$$

If we *do* know what is in the box, we can determine $\mathcal{E}_{eq}$ and R_{eq} by calculating them instead of measuring them.

- For $\mathcal{E}_{eq}$, calculate the open circuit voltage between the two terminals (with nothing connected to them outside the box). In the above example, this is just I_3R_3, with I_3 given by Eq. (4.33).
- For R_{eq}, connect the terminals by a wire with zero resistance, and calculate the short-circuit current I_{sc} through this wire; R_{eq} is then given by $\mathcal{E}_{eq}/I_{sc}$. See Problem 4.14 for how this works in the above example. There is, however, a second method for calculating R_{eq}, which is generally much quicker: R_{eq} is the resistance that would be measured between the two terminals with all the internal electromotive forces made zero. In our example that would be the resistance of R_1, R_2, and R_3 all in parallel, which is $R_1R_2R_3/(R_1R_2 + R_2R_3 + R_1R_3)$. The reason why this method works is explained in the solution to Problem 4.13.

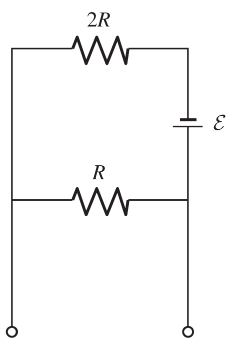

Figure 4.25.
Find $\mathcal{E}_{eq}$ and R_{eq} for this circuit.

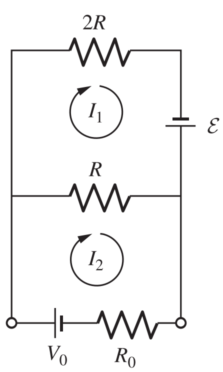

Figure 4.26.
Loop currents for use in Kirchhoff’s rules.

Example

(a) Find the Thévenin equivalent $\mathcal{E}_{eq}$ and R_{eq} for the circuit shown in Fig. 4.25.
(b) Calculate $\mathcal{E}_{eq}$ and R_{eq} again, but now do it the long way. Use Kirchhoff’s rules to find the current passing through the bottom branch of the circuit in Fig. 4.26, and then interpret your result in a way that gives you $\mathcal{E}_{eq}$ and R_{eq}.

Solution

(a) $\mathcal{E}_{eq}$ is the open-circuit voltage. With nothing connected to the terminals, the current running around the loop is $\mathcal{E}/3R$. The voltage drop across the R resistor is therefore $(\mathcal{E}/3R)(R) = \mathcal{E}/3$. But this is also the open-circuit voltage between the two terminals, so $\mathcal{E}_{eq} = \mathcal{E}/3$.

We can find R_{eq} in two ways. The quick way is to calculate the resistance between the terminals with $\mathcal{E}$ set equal to zero. In that case we have an R and a $2R$ in parallel, so $R_{eq} = 2R/3$.

Alternatively, we can find R_{eq} by calculating the short-circuit current between the terminals. With the short circuit present, no current takes the route through the R resistor, so we just have $\mathcal{E}$ and $2R$ in series. The short-circuit current between the terminals is therefore $I_{sc} = \mathcal{E}/2R$. The equivalent resistance is then given by $R_{eq} = \mathcal{E}_{eq}/I_{sc} = (\mathcal{E}/3)/(\mathcal{E}/2R) = 2R/3$.

(b) The loop equations for the circuit in Fig. 4.26 are

$$0 = \mathcal{E} - R(I_1 - I_2) - (2R)I_1,$$
$$0 = V_0 - R(I_2 - I_1) - R_0 I_2. \tag{4.36}$$

Solving these equations for I_2 gives $I_2 = (\mathcal{E} + 3V_0)/(2R + 3R_0)$ (as you can check), which can be written suggestively as

$$V_0 + \frac{\mathcal{E}}{3} = I_2\left(R_0 + \frac{2R}{3}\right). \tag{4.37}$$

But this is exactly the $V = IR$ statement that we would write down for the circuit shown in Fig. 4.27, where the total emf is $V_0 + \mathcal{E}/3$ and the total resistance is $R_0 + 2R/3$. Since the result in Eq. (4.37) holds for any values of V_0 and R_0, we conclude that the given circuit is equivalent to an emf $\mathcal{E}_{eq} = \mathcal{E}/3$ in series with a resistor $R_{eq} = 2R/3$. Generalizing this method is the basic idea behind the first proof of Thévenin's theorem given in Problem 4.13.

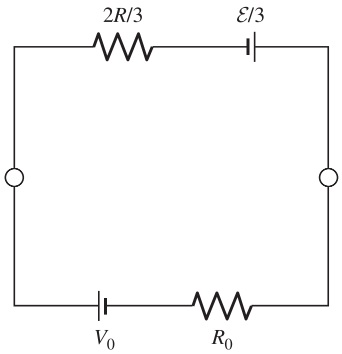

Figure 4.27.
The Thévenin equivalent circuit.

In analyzing a complicated circuit it sometimes helps to replace a two-terminal section by its equivalent $\mathcal{E}_{eq}$ and R_{eq}. Thévenin's theorem assumes the linearity of all circuit elements, including the reversibility of currents through batteries. If one of our batteries is a nonrechargeable dry cell with the current through it backward, caution is advisable!

4.11 Variable currents in capacitors and resistors

Let a capacitor of capacitance C be charged to some potential V_0 and then discharged by suddenly connecting it across a resistance R. Figure 4.28 shows the capacitor indicated by the conventional symbol , the resistor R, and a switch which we shall imagine to be closed at time $t = 0$. It is obvious that, as current flows, the capacitor will gradually lose its charge, the voltage across the capacitor will diminish, and this in turn will lessen the flow of current. Let's be quantitative about this.

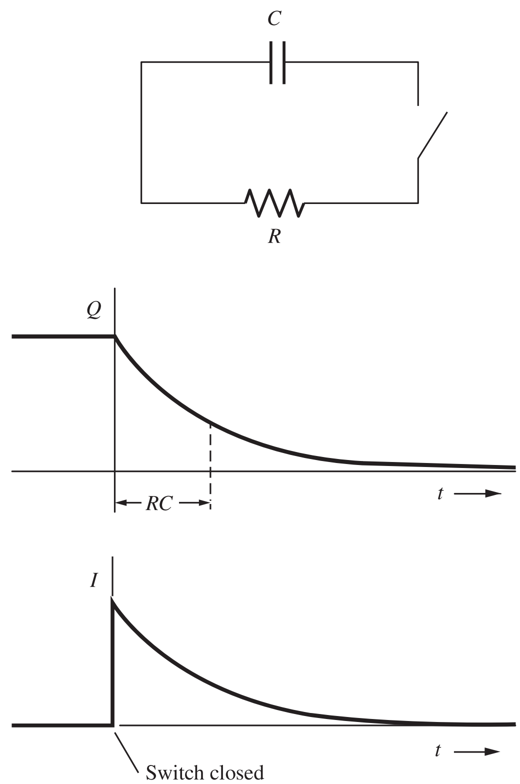

Figure 4.28.
Charge and current in an RC circuit. Both quantities decay by the factor $1/e$ in time RC.

Example (*RC* circuit) In the circuit in Fig. 4.28, what are the charge Q on the capacitor and the current I in the circuit, as functions of time?

Solution To find $Q(t)$ and $I(t)$ we need only write down the conditions that govern the circuit. Let $V(t)$ be the potential difference between the plates, which is also the voltage across the resistor R. Let the current I be considered positive if it flows away from the positive side of the capacitor. The quantities Q, I, and V, all functions of the time, must be related as follows:

$$Q = CV, \qquad I = \frac{V}{R}, \qquad -\frac{dQ}{dt} = I. \tag{4.38}$$

Eliminating I and V, we obtain the equation that governs the time variation of Q:

$$\frac{dQ}{dt} = -\frac{Q}{RC}. \tag{4.39}$$

Writing this in the form

$$\frac{dQ}{Q} = -\frac{dt}{RC}, \tag{4.40}$$

we can integrate both sides, obtaining

$$\ln Q = \frac{-t}{RC} + \text{const.} \tag{4.41}$$

The solution of our differential equation is therefore

$$Q = (\text{another constant}) \cdot e^{-t/RC}. \tag{4.42}$$

If $V = V_0$ at $t = 0$, then $Q = CV_0$ at $t = 0$. This determines the constant, and we now have the exact behavior of Q after the switch is closed:

$$Q(t) = CV_0\, e^{-t/RC}. \tag{4.43}$$

The behavior of the current I is found directly from this:

$$I(t) = -\frac{dQ}{dt} = \frac{V_0}{R}\, e^{-t/RC}. \tag{4.44}$$

And the voltage at any time is $V(t) = I(t)R$, or alternatively $V(t) = Q(t)/C$.

At the closing of the switch the current rises at once to the value V_0/R and then decays exponentially to zero. The time that characterizes this decay is the constant RC in the above exponents. People often speak of the "RC time constant" associated with a circuit or part of a circuit. Let's double check that RC does indeed have units of time. In SI units, R is measured in ohms, which from Eq. (4.18) is given by volt/ampere. And C is measured in farads, which from Eq. (3.8) is given by coulomb/volt. So RC has units of coulomb/ampere, which is a second, as desired. If we make the circuit in Fig. 4.28 out of a 0.05 microfarad capacitor and a 5 megohm resistor, both of which are reasonable objects to find around any laboratory, we would have $RC = (5 \cdot 10^6\ \text{ohm})(0.05 \cdot 10^{-6}\ \text{farad}) = 0.25\ \text{s}$.

Quite generally, in any electrical system made up of charged conductors and resistive current paths, one time scale – perhaps not the only one – for processes in the system is set by some resistance–capacitance product. This has a bearing on our earlier observation on page 187 that $\epsilon_0\rho$ has the dimensions of time. Imagine a capacitor with plates of area A and separation s. Its capacitance C is $\epsilon_0 A/s$. Now imagine the space between the plates suddenly filled with a conductive medium of resistivity ρ. To avoid any question of how this might affect the capacitance, let us suppose that the medium is a very slightly ionized gas; a substance of that density will hardly affect the capacitance at all. This new conductive path will discharge the capacitor as effectively as did the external resistor in Fig. 4.28. How quickly will this happen? From Eq. (4.17) the resistance of the path, R, is $\rho s/A$. Hence the time constant RC is just $(\rho s/A)(\epsilon_0 A/s) = \epsilon_0\rho$. For example, if our weakly ionized gas had a resistivity of 10^6 ohm-meter, the time constant for discharge of the capacitor would be (recalling the units of ϵ_0 and the ohm) $\epsilon_0\rho = (8.85 \cdot 10^{-12}\,\text{C}^2\,\text{s}^2\,\text{kg}^{-1}\,\text{m}^{-3})(10^6\,\text{kg}\,\text{m}^3\,\text{C}^{-2}\,\text{s}^{-1}) \approx 10$ microseconds. It does not depend on the size or shape of the capacitor.

What we have here is simply the time constant for the relaxation of an electric field in a conducting medium by redistribution of charge. We really don't need the capacitor plates to describe it. Imagine that we could suddenly imbed two sheets of charge, a negative sheet and a positive sheet, opposite one another in a conductor – for instance, in an n-type semiconductor (Fig. 4.29(a)). What will make these charges disappear? Do negative charge carriers move from the sheet on the left across the intervening space, neutralizing the positive charges when they arrive at the sheet on the right? Surely not – if that were the process, the time required would be proportional to the distance between the sheets. What happens instead is this. The *entire population* of negative charge carriers that fills the space between the sheets is caused to move by the electric field. Only a *very slight* displacement of this cloud of charge suffices to remove excess negative charge on the left, while providing on the right the extra negative charge needed to neutralize the positive sheet, as indicated in Fig. 4.29(b). Within a conductor, in other words, neutrality is restored by a small readjustment of the entire charge distribution, not by a few charge carriers moving a long distance. That is why the relaxation time can be independent of the size of the system.

For a metal with resistivity typically 10^{-7} ohm-meter, the time constant $\epsilon_0\rho$ is about 10^{-18} s, orders of magnitude shorter than the mean free time of a conduction electron in the metal. As a relaxation time this makes no sense. Our theory, at this stage, can tell us nothing about events on a time scale as short as that.

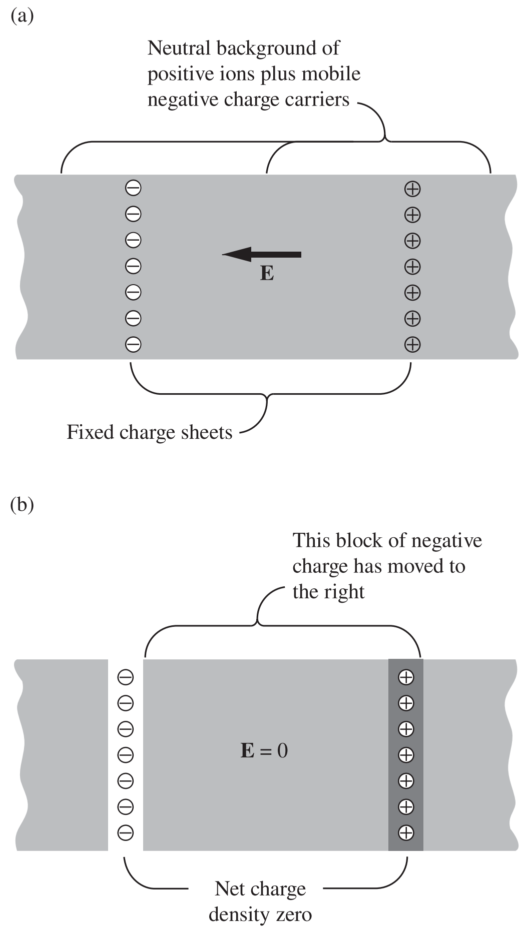

Figure 4.29.
In a conducting medium, here represented by an n-type conductor, two fixed sheets of charge, one negative and one positive, can be neutralized by a slight motion of the entire block of mobile charge carriers lying between them. (a) Before the block of negative charge has moved. (b) After the net charge density has been reduced to zero at each sheet.

4.12 Applications

The *transatlantic telegraph cable* (see Exercise 4.22) extended about 2000 miles between Newfoundland and Ireland, and was the most expensive and involved electrical engineering project of its time. After many failures, interrupted by a very short-lived success in 1858, it was finally completed in 1866. The initial failures were due partly to the fact that there didn't exist a consistent set of electrical units, in particular a unit