1

Electrostatics: charges and fields

Overview The existence of this book is owed (both figuratively and literally) to the fact that the building blocks of matter possess a quality called *charge.* Two important aspects of charge are *conservation* and *quantization.* The electric force between two charges is given by *Coulomb's law.* Like the gravitational force, the electric force falls off like $1/r^2$. It is *conservative,* so we can talk about the potential energy of a system of charges (the work done in assembling them). A very useful concept is the *electric field,* which is defined as the force per unit charge. Every point in space has a unique electric field associated with it. We can define the *flux* of the electric field through a given surface. This leads us to *Gauss's law,* which is an alternative way of stating Coulomb's law. In cases involving sufficient symmetry, it is much quicker to calculate the electric field via Gauss's law than via Coulomb's law and direct integration. Finally, we discuss the *energy density* in the electric field, which provides another way of calculating the potential energy of a system.

1.1 Electric charge

Electricity appeared to its early investigators as an extraordinary phenomenon. To draw from bodies the "subtle fire," as it was sometimes called, to bring an object into a highly electrified state, to produce a steady flow of current, called for skillful contrivance. Except for the spectacle of lightning, the ordinary manifestations of nature, from the freezing of water to the growth of a tree, seemed to have no relation to the curious behavior of electrified objects. We know now that electrical forces largely determine the physical and chemical properties of matter over the whole range from atom to living cell. For this understanding we have to thank the scientists of the nineteenth century, Ampère, Faraday, Maxwell, and many others, who discovered the nature of electromagnetism, as well as the physicists and chemists of the twentieth century who unraveled the atomic structure of matter.

Classical electromagnetism deals with electric charges and currents and their interactions as if all the quantities involved could be measured independently, with unlimited precision. Here *classical* means simply "nonquantum." The quantum law with its constant h is ignored in the classical theory of electromagnetism, just as it is in ordinary mechanics. Indeed, the classical theory was brought very nearly to its present state of completion before Planck's discovery of quantum effects in 1900. It has survived remarkably well. Neither the revolution of quantum physics nor the development of special relativity dimmed the luster of the electromagnetic field equations Maxwell wrote down 150 years ago.

Of course the theory was solidly based on experiment, and because of that was fairly secure within its original range of application – to coils, capacitors, oscillating currents, and eventually radio waves and light waves. But even so great a success does not guarantee validity in another domain, for instance, the inside of a molecule.

Two facts help to explain the continuing importance in modern physics of the classical description of electromagnetism. First, special relativity required no revision of classical electromagnetism. Historically speaking, special relativity *grew out* of classical electromagnetic theory and experiments inspired by it. Maxwell's field equations, developed long before the work of Lorentz and Einstein, proved to be entirely compatible with relativity. Second, quantum modifications of the electromagnetic forces have turned out to be unimportant down to distances less than 10^{-12} meters, 100 times smaller than the atom. We can describe the repulsion and attraction of particles in the atom using the same laws that apply to the leaves of an electroscope, although we need quantum mechanics to predict how the particles will behave under those forces. For still smaller distances, a fusion of electromagnetic theory and quantum theory, called *quantum electrodynamics,* has been remarkably successful. Its predictions are confirmed by experiment down to the smallest distances yet explored.

It is assumed that the reader has some acquaintance with the elementary facts of electricity. We are not going to review all the experiments by which the existence of electric charge was demonstrated, nor shall we review all the evidence for the electrical constitution of matter. On the other hand, we do want to look carefully at the experimental foundations of the basic laws on which all else depends. In this chapter we shall study the physics of stationary electric charges – *electrostatics.*

Certainly one fundamental property of electric charge is its existence in the two varieties that were long ago named *positive* and *negative.*

The observed fact is that all charged particles can be divided into two classes such that all members of one class repel each other, while attracting members of the other class. If two small electrically charged bodies A and B, some distance apart, attract one another, and if A attracts some third electrified body C, then we always find that B repels C. Contrast this with gravitation: there is only one kind of gravitational mass, and every mass attracts every other mass.

One may regard the two kinds of charge, positive and negative, as opposite manifestations of one quality, much as *right* and *left* are the two kinds of handedness. Indeed, in the physics of elementary particles, questions involving the sign of the charge are sometimes linked to a question of handedness, and to another basic symmetry, the relation of a sequence of events, a, then b, then c, to the temporally reversed sequence c, then b, then a. It is only the duality of electric charge that concerns us here. For every kind of particle in nature, as far as we know, there can exist an *antiparticle,* a sort of electrical "mirror image." The antiparticle carries charge of the opposite sign. If any other intrinsic quality of the particle has an opposite, the antiparticle has that too, whereas in a property that admits no opposite, such as mass, the antiparticle and particle are exactly alike.

The electron's charge is negative; its antiparticle, called a *positron,* has a positive charge, but its mass is precisely the same as that of the electron. The proton's antiparticle is called simply an *antiproton*; its electric charge is negative. An electron and a proton combine to make an ordinary hydrogen atom. A positron and an antiproton could combine in the same way to make an atom of antihydrogen. Given the building blocks, positrons, antiprotons, and antineutrons,[1] there could be built up the whole range of antimatter, from antihydrogen to antigalaxies. There is a practical difficulty, of course. Should a positron meet an electron or an antiproton meet a proton, that pair of particles will quickly vanish in a burst of radiation. It is therefore not surprising that even positrons and antiprotons, not to speak of antiatoms, are exceedingly rare and short-lived in our world. Perhaps the universe contains, somewhere, a vast concentration of antimatter. If so, its whereabouts is a cosmological mystery.

The universe around us consists overwhelmingly of matter, not antimatter. That is to say, the abundant carriers of negative charge are electrons, and the abundant carriers of positive charge are protons. The proton is nearly 2000 times heavier than the electron, and very different, too, in some other respects. Thus matter at the atomic level incorporates negative and positive electricity in quite different ways. The positive charge is all in the atomic nucleus, bound within a massive structure no more than 10^{-14} m in size, while the negative charge is spread, in

[1] Although the electric charge of each is zero, the neutron and its antiparticle are not interchangeable. In certain properties that do not concern us here, they are opposite.

effect, through a region about 10^4 times larger in dimensions. It is hard to imagine what atoms and molecules – and all of chemistry – would be like, if not for this fundamental electrical asymmetry of matter.

What we call negative charge, by the way, could just as well have been called positive. The name was a historical accident. There is nothing essentially negative about the charge of an electron. It is not like a negative integer. A negative integer, once multiplication has been defined, differs essentially from a positive integer in that its square is an integer of opposite sign. But the product of two charges is not a charge; there is no comparison.

Two other properties of electric charge are essential in the electrical structure of matter: charge is *conserved,* and charge is *quantized.* These properties involve *quantity* of charge and thus imply a measurement of charge. Presently we shall state precisely how charge can be measured in terms of the force between charges a certain distance apart, and so on. But let us take this for granted for the time being, so that we may talk freely about these fundamental facts.

1.2 Conservation of charge

The total charge in an isolated system never changes. By *isolated* we mean that no matter is allowed to cross the boundary of the system. We could let light pass into or out of the system, since the "particles" of light, called *photons,* carry no charge at all. Within the system charged particles may vanish or reappear, but they always do so in pairs of equal and opposite charge. For instance, a thin-walled box in a vacuum exposed to gamma rays might become the scene of a "pair-creation" event in which a high-energy photon ends its existence with the creation of an electron and a positron (Fig. 1.1). Two electrically charged particles have been newly created, but the net change in total charge, in and on the box, is zero. An event that *would* violate the law we have just stated would be the creation of a positively charged particle *without* the simultaneous creation of a negatively charged particle. Such an occurrence has never been observed.

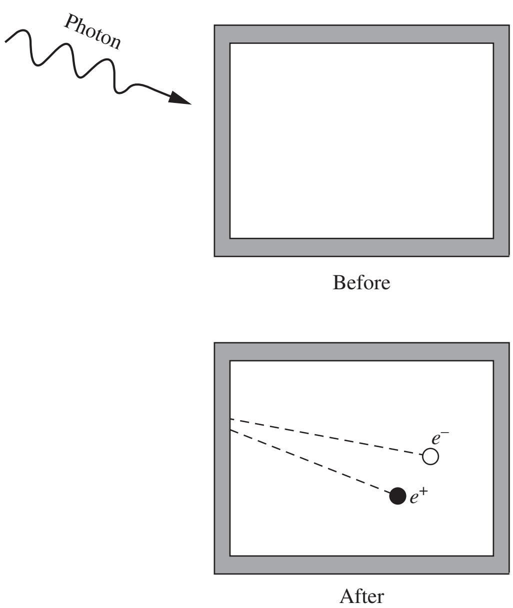

Figure 1.1.
Charged particles are created in pairs with equal and opposite charge.

Of course, if the electric charges of an electron and a positron were not precisely equal in magnitude, pair creation would still violate the strict law of charge conservation. That equality is a manifestation of the particle–antiparticle duality already mentioned, a universal symmetry of nature.

One thing will become clear in the course of our study of electromagnetism: nonconservation of charge would be quite incompatible with the structure of our present electromagnetic theory. We may therefore state, either as a postulate of the theory or as an empirical law supported without exception by all observations so far, the charge conservation law:

The total electric charge in an isolated system, that is, the algebraic sum of the positive and negative charge present at any time, never changes.

Sooner or later we must ask whether this law meets the test of relativistic invariance. We shall postpone until Chapter 5 a thorough discussion of this important question. But the answer is that it does, and not merely in the sense that the statement above holds in any given inertial frame, but in the stronger sense that observers in different frames, measuring the charge, obtain the same number. In other words, the total electric charge of an isolated system is a relativistically invariant number.

1.3 Quantization of charge

The electric charges we find in nature come in units of one magnitude only, equal to the amount of charge carried by a single electron. We denote the magnitude of that charge by e. (When we are paying attention to sign, we write $-e$ for the charge on the electron itself.) We have already noted that the positron carries precisely that amount of charge, as it must if charge is to be conserved when an electron and a positron annihilate, leaving nothing but light. What seems more remarkable is the apparently exact equality of the charges carried by all other charged particles – the equality, for instance, of the positive charge on the proton and the negative charge on the electron.

That particular equality is easy to test experimentally. We can see whether the net electric charge carried by a hydrogen molecule, which consists of two protons and two electrons, is zero. In an experiment carried out by J. G. King,[2] hydrogen gas was compressed into a tank that was electrically insulated from its surroundings. The tank contained about $5 \cdot 10^{24}$ molecules (approximately 17 grams) of hydrogen. The gas was then allowed to escape by means that prevented the escape of any ion – a molecule with an electron missing or an extra electron attached. If the charge on the proton differed from that on the electron by, say, one part in a billion, then each hydrogen molecule would carry a charge of $2 \cdot 10^{-9}e$, and the departure of the whole mass of hydrogen would alter the charge of the tank by $10^{16}e$, a gigantic effect. In fact, the experiment could have revealed a residual molecular charge as small as $2 \cdot 10^{-20}e$, and none was observed. This proved that the proton and the electron do not differ in magnitude of charge by more than 1 part in 10^{20}.

Perhaps the equality is really *exact* for some reason we don't yet understand. It may be connected with the possibility, suggested by certain

[2] See King (1960). References to previous tests of charge equality will be found in this article and in the chapter by V. W. Hughes in Hughes (1964).

theories, that a proton can, *very* rarely, decay into a positron and some uncharged particles. If that were to occur, even the slightest discrepancy between proton charge and positron charge would violate charge conservation. Several experiments designed to detect the decay of a proton have not yet, as of this writing, registered with certainty a single decay. If and when such an event is observed, it will show that exact equality of the magnitude of the charge of the proton and the charge of the electron (the positron's antiparticle) can be regarded as a corollary of the more general law of charge conservation.

That notwithstanding, we now know that the *internal* structure of all the strongly interacting particles called *hadrons* – a class that includes the proton and the neutron – involves basic units called *quarks,* whose electric charges come in multiples of $e/3$. The proton, for example, is made with three quarks, two with charge $2e/3$ and one with charge $-e/3$. The neutron contains one quark with charge $2e/3$ and two quarks with charge $-e/3$.

Several experimenters have searched for single quarks, either free or attached to ordinary matter. The fractional charge of such a quark, since it cannot be neutralized by any number of electrons or protons, should betray the quark's presence. So far no fractionally charged particle has been conclusively identified. The present theory of the strong interactions, called *quantum chromodynamics,* explains why the liberation of a quark from a hadron is most likely impossible.

The fact of charge quantization lies outside the scope of classical electromagnetism, of course. We shall usually ignore it and act as if our point charges q could have any strength whatsoever. This will not get us into trouble. Still, it is worth remembering that classical theory cannot be expected to explain the structure of the elementary particles. (It is not certain that present quantum theory can either!) What holds the electron together is as mysterious as what fixes the precise value of its charge. Something more than electrical forces must be involved, for the electrostatic forces between different parts of the electron would be repulsive.

In our study of electricity and magnetism we shall treat the charged particles simply as carriers of charge, with dimensions so small that their extension and structure is, for most purposes, quite insignificant. In the case of the proton, for example, we know from high-energy scattering experiments that the electric charge does not extend appreciably beyond a radius of 10^{-15} m. We recall that Rutherford's analysis of the scattering of alpha particles showed that even heavy nuclei have their electric charge distributed over a region smaller than 10^{-13} m. For the physicist of the nineteenth century a "point charge" remained an abstract notion. Today we are on familiar terms with the atomic particles. The graininess of electricity is so conspicuous in our modern description of nature that we find a point charge less of an artificial idealization than a smoothly varying distribution of charge density. When we postulate such smooth charge distributions, we may think of them as averages over very large numbers of elementary charges, in the same way that we can define the macroscopic density of a liquid, its lumpiness on a molecular scale notwithstanding.

1.4 Coulomb's law

As you probably already know, the interaction between electric charges at rest is described by Coulomb's law: two stationary electric charges repel or attract one another with a force proportional to the product of the magnitude of the charges and inversely proportional to the square of the distance between them.

We can state this compactly in vector form:

$$\mathbf{F}_2 = k\frac{q_1 q_2 \hat{\mathbf{r}}_{21}}{r_{21}^2}. \tag{1.1}$$

Here q_1 and q_2 are numbers (scalars) giving the magnitude and sign of the respective charges, $\hat{\mathbf{r}}_{21}$ is the unit vector in the direction[3] from charge 1 to charge 2, and $\mathbf{F}_2$ is the force acting on charge 2. Thus Eq. (1.1) expresses, among other things, the fact that like charges repel and unlike charges attract. Also, the force obeys Newton's third law; that is, $\mathbf{F}_2 = -\mathbf{F}_1$.

The unit vector $\hat{\mathbf{r}}_{21}$ shows that the force is parallel to the line joining the charges. It could not be otherwise unless space itself has some built-in directional property, for with two point charges alone in empty and isotropic space, no other direction could be singled out.

If the point charge itself had some internal structure, with an axis defining a direction, then it would have to be described by more than the mere scalar quantity q. It is true that some elementary particles, including the electron, do have another property, called *spin*. This gives rise to a magnetic force between two electrons in addition to their electrostatic repulsion. This magnetic force does not, in general, act in the direction of the line joining the two particles. It decreases with the inverse fourth power of the distance, and at atomic distances of 10^{-10} m the Coulomb force is already about 10^4 times stronger than the magnetic interaction of the spins. Another magnetic force appears if our charges are moving – hence the restriction to stationary charges in our statement of Coulomb's law. We shall return to these magnetic phenomena in later chapters.

Of course we must assume, in writing Eq. (1.1), that both charges are well localized, each occupying a region small compared with r_{21}. Otherwise we could not even define the distance r_{21} precisely.

The value of the constant k in Eq. (1.1) depends on the units in which r, $\mathbf{F}$, and q are to be expressed. In this book we will use the International System of Units, or "SI" units for short. This system is based on the

[3] The convention we adopt here may not seem the natural choice, but it is more consistent with the usage in some other parts of physics and we shall try to follow it throughout this book.

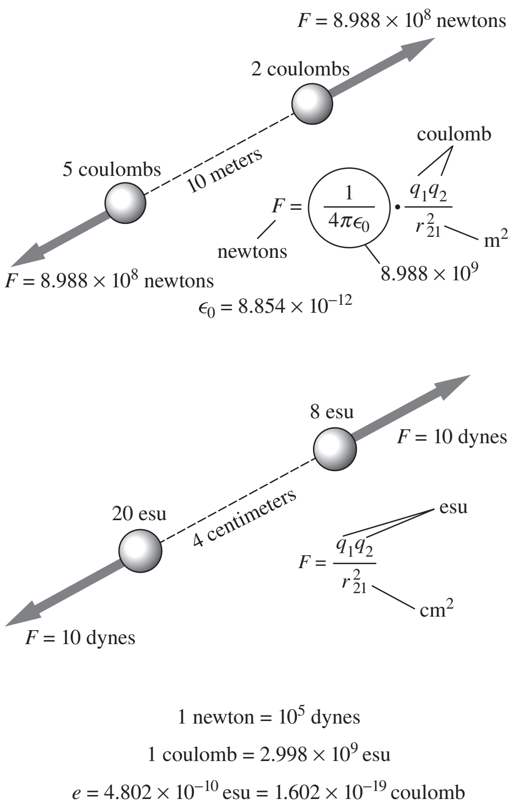

Figure 1.2.
Coulomb's law expressed in SI units (top) and in Gaussian electrostatic units (bottom). The constant ϵ_0 and the factor relating coulombs to esu are connected, as we shall learn later, with the speed of light. We have rounded off the constants in the figure to four-digit accuracy. The precise values are given in Appendix E.

meter, kilogram, and second as units of length, mass, and time. The SI unit of charge is the *coulomb* (C). Some other SI electrical units that we will eventually become familiar with are the volt, ohm, ampere, and tesla. The official definition of the coulomb involves the magnetic force, which we will discuss in Chapter 6. For present purposes, we can define the coulomb as follows. Two like charges, each of 1 coulomb, repel one another with a force of $8.988 \cdot 10^9$ newtons when they are 1 meter apart. In other words, the k in Eq. (1.1) is given by

$$k = 8.988 \cdot 10^9 \,\frac{\text{N m}^2}{\text{C}^2}. \tag{1.2}$$

In Chapter 6 we will learn where this seemingly arbitrary value of k comes from. In general, approximating k by $9 \cdot 10^9\ \text{N m}^2/\text{C}^2$ is quite sufficient. The magnitude of e, the fundamental quantum of electric charge, happens to be about $1.602 \cdot 10^{-19}$ C. So if you wish, you may think of a coulomb as defined to be the magnitude of the charge contained in $6.242 \cdot 10^{18}$ electrons.

Instead of k, it is customary (for historical reasons) to introduce a constant ϵ_0 which is defined by

$$k \equiv \frac{1}{4\pi\epsilon_0} \quad\Longrightarrow\quad \boxed{\epsilon_0 \equiv \frac{1}{4\pi k} = 8.854 \cdot 10^{-12}\,\frac{\text{C}^2}{\text{N m}^2}} \quad \left(\text{or } \frac{\text{C}^2\,\text{s}^2}{\text{kg m}^3}\right). \tag{1.3}$$

In terms of ϵ_0, Coulomb's law in Eq. (1.1) takes the form

$$\boxed{\mathbf{F} = \frac{1}{4\pi\epsilon_0}\,\frac{q_1 q_2 \hat{\mathbf{r}}_{21}}{r_{21}^2}} \tag{1.4}$$

The constant ϵ_0 will appear in many expressions that we will meet in the course of our study. The 4π is included in the definition of ϵ_0 so that certain formulas (such as Gauss's law in Sections 1.10 and 2.9) take on simple forms. Additional details and technicalities concerning ϵ_0 can be found in Appendix E.

Another system of units that comes up occasionally is the *Gaussian* system, which is one of several types of cgs systems, short for centimeter–gram–second. (In contrast, the SI system is an mks system, short for meter–kilogram–second.) The Gaussian unit of charge is the "electrostatic unit," or esu. The esu is defined so that the constant k in Eq. (1.1) *exactly* equals 1 (and this is simply the number 1, with no units) when r_{21} is measured in cm, $\mathbf{F}$ in dynes, and the q values in esu. Figure 1.2 gives some examples using the SI and Gaussian systems of units. Further discussion of the SI and Gaussian systems can be found in Appendix A.

Example (Relation between 1 coulomb and 1 esu) Show that 1 coulomb equals $2.998 \cdot 10^9$ esu (which generally can be approximated by $3 \cdot 10^9$ esu).

Solution From Eqs. (1.1) and (1.2), two charges of 1 coulomb separated by a distance of 1 m exert a (large!) force of $8.988 \cdot 10^9\ \text{N} \approx 9 \cdot 10^9\ \text{N}$ on each other. We can convert this to the Gaussian unit of force via

$$1\,\text{N} = 1\,\frac{\text{kg m}}{\text{s}^2} = \frac{(1000\,\text{g})(100\,\text{cm})}{\text{s}^2} = 10^5\,\frac{\text{g cm}}{\text{s}^2} = 10^5\ \text{dynes}. \tag{1.5}$$

The two 1 C charges therefore exert a force of $9 \cdot 10^{14}$ dynes on each other. How would someone working in Gaussian units describe this situation? In Gaussian units, Coulomb's law gives the force simply as q^2/r^2. The separation is 100 cm, so if 1 coulomb equals N esu (with N to be determined), the $9 \cdot 10^{14}$ dyne force between the charges can be expressed as

$$9 \cdot 10^{14}\ \text{dyne} = \frac{(N\,\text{esu})^2}{(100\,\text{cm})^2} \implies N^2 = 9 \cdot 10^{18} \implies N = 3 \cdot 10^9. \tag{1.6}$$

Hence,[4]

$$1\,\text{C} = 3 \cdot 10^9\ \text{esu}. \tag{1.7}$$

The magnitude of the electron charge is then given approximately by $e = 1.6 \cdot 10^{-19}\,\text{C} \approx 4.8 \cdot 10^{-10}$ esu.

If we had used the more exact value of k in Eq. (1.2), the "3" in our result would have been replaced by $\sqrt{8.988} = 2.998$. This looks suspiciously similar to the "2.998" in the speed of light, $c = 2.998 \cdot 10^8$ m/s. This is no coincidence. We will see in Section 6.1 that Eq. (1.7) can actually be written as $1\,\text{C} = (10\{c\})$ esu, where we have put the c in brackets to signify that it is just the number $2.998 \cdot 10^8$ without the units of m/s.

On an everyday scale, a coulomb is an extremely large amount of charge, as evidenced by the fact that if you have two such charges separated by 1 m (never mind how you would keep each charge from flying apart due to the self repulsion!), the above force of $9 \cdot 10^9$ N between them is about one million tons. The esu is a much more reasonable unit to use for everyday charges. For example, the static charge on a balloon that sticks to your hair is on the order of 10 or 100 esu.

The only way we have of detecting and measuring electric charges is by observing the interaction of charged bodies. One might wonder, then, how much of the apparent content of Coulomb's law is really only definition. As it stands, the significant physical content is the statement of inverse-square dependence and the implication that electric charge

[4] We technically shouldn't be using an "=" sign here, because it suggests that the units of a coulomb are the same as those of an esu. This is not the case; they are units in different systems and cannot be expressed in terms of each other. The proper way to express Eq. (1.7) is to say, "1 C is equivalent to $3 \cdot 10^9$ esu." But we'll usually just use the "=" sign, and you'll know what we mean. See Appendix A for further discussion of this.

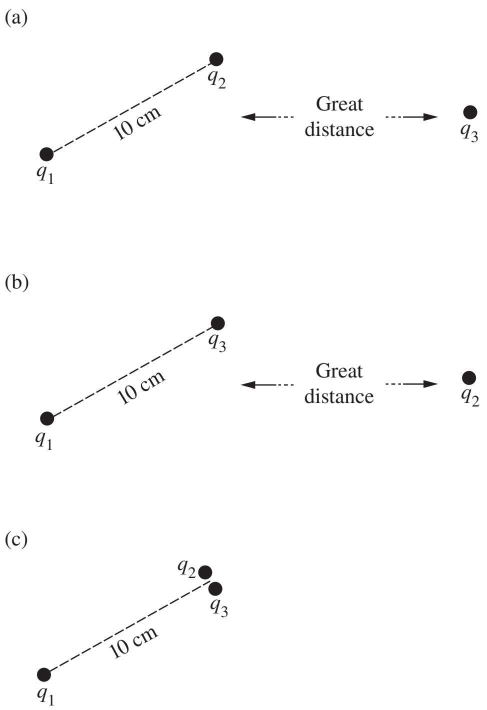

Figure 1.3.
The force on q_1 in (c) is the sum of the forces on q_1 in (a) and (b).

is *additive* in its effect. To bring out the latter point, we have to consider *more* than two charges. After all, if we had only two charges in the world to experiment with, q_1 and q_2, we could never measure them separately. We could verify only that F is proportional to $1/r_{21}^2$. Suppose we have *three* bodies carrying charges q_1, q_2, and q_3. We can measure the force on q_1 when q_2 is 10 cm away from q_1, with q_3 very far away, as in Fig. 1.3(a). Then we can take q_2 away, bring q_3 into q_2's former position, and again measure the force on q_1. Finally, we can bring q_2 and q_3 very close together and locate the combination 10 cm from q_1. We find by measurement that the force on q_1 is equal to the sum of the forces previously measured. This is a significant result that could *not* have been predicted by logical arguments from symmetry like the one we used above to show that the force between two point charges *had* to be along the line joining them. *The force with which two charges interact is not changed by the presence of a third charge.*

No matter how many charges we have in our system, Coulomb's law in Eq. (1.4) can be used to calculate the interaction of every pair. This is the basis of the principle of *superposition,* which we shall invoke again and again in our study of electromagnetism. Superposition means combining two sets of sources into one system by adding the second system "on top of" the first without altering the configuration of either one. Our principle ensures that the force on a charge placed at any point in the combined system will be the vector sum of the forces that each set of sources, acting alone, causes to act on a charge at that point. This principle must not be taken lightly for granted. There may well be a domain of phenomena, involving very small distances or very intense forces, where superposition *no longer holds.* Indeed, we know of quantum phenomena in the electromagnetic field that do represent a failure of superposition, seen from the viewpoint of the classical theory.

Thus the physics of electrical interactions comes into full view only when we have *more* than two charges. We can go beyond the explicit statement of Eq. (1.1) and assert that, with the three charges in Fig. 1.3 occupying any positions whatsoever, the force on any one of them, such as q_3, is correctly given by the following equation:

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{q_3 q_1 \hat{\mathbf{r}}_{31}}{r_{31}^2} + \frac{1}{4\pi\epsilon_0}\frac{q_3 q_2 \hat{\mathbf{r}}_{32}}{r_{32}^2}. \tag{1.8}$$

The experimental verification of the inverse-square law of electrical attraction and repulsion has a curious history. Coulomb himself announced the law in 1786 after measuring with a torsion balance the force between small charged spheres. But 20 years earlier Joseph Priestly, carrying out an experiment suggested to him by Benjamin Franklin, had noticed the absence of electrical influence within a hollow charged container and made an inspired conjecture: "May we not infer from this experiment that the attraction of electricity is subject to the same laws with that of gravitation and is therefore according to the square of the